

THIS NUMBER CONTAINS THE INDEX FOR THE ENTIRE VOLUME

VOLUME 57
NUMBER 5

WHOLE No. 265
1944

Psychological Monographs

JOHN F. DASHIELL, *Editor*

An Experimental Application of Projective Principles to a Paper and Pencil Personality Test

By

HELEN SARGENT
Northwestern University

Published by

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

Publications Office

NORTHWESTERN UNIVERSITY, EVANSTON, ILLINOIS

Collected set

PUBLICATIONS OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION

WILLARD L. VALENTINE, *Business Manager*

PSYCHOLOGICAL REVIEW

HERBERT S. LANGFELD, *Editor*
Princeton University

Contains original contributions only, appears bi-monthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 540 pages.

Subscription: \$5.50 (Foreign, \$5.75). Single copies, \$1.00.

PSYCHOLOGICAL BULLETIN

JOHN E. ANDERSON, *Editor*
University of Minnesota

Contains critical reviews of books and articles, psychological news and notes, university notices, and announcements. Appears monthly (10 issues), the annual volume comprising about 665 pages. Special issue of the BULLETIN consist of general reviews of recent work in some department of psychology.

Subscription: \$7.00 (Foreign, \$7.25). Single copies, 75¢.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

S. W. FERNBERGER, *Editor*
University of Pennsylvania

Contains original contributions of an experimental character. Appears bi-monthly, one volume per year.

Subscription: \$7.00 per volume; Foreign, \$7.25. Single copies, \$1.25.

PSYCHOLOGICAL ABSTRACTS

WALTER S. HUNTER, *Editor*
Brown University

Appears monthly, the twelve numbers and an index supplement making a volume of about 700 pages. The journal is devoted to the publication of non-critical abstracts of the world's literature in psychology and closely related subjects.

Subscription: \$7.00 (Foreign, \$7.25). Single copies, 75¢.

PSYCHOLOGICAL MONOGRAPHS

JOHN F. DASHIELL, *Editor*
University of North Carolina

Consists of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 350 pages. Two volumes a year.

Subscription: \$6.00 per volume (Foreign, \$6.30).

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

GORDON W. ALLPORT, *Editor*
Harvard University

Appears quarterly, January, April, July, October, the four numbers comprising a volume of 400 pages. The journal contains original contributions in the field of abnormal and social psychology, reviews, and notes.

Subscription: \$5.00 (Foreign, \$5.25). Single copies, \$1.50.

JOURNAL OF APPLIED PSYCHOLOGY

DONALD G. PATERSON, *Editor*
University of Minnesota

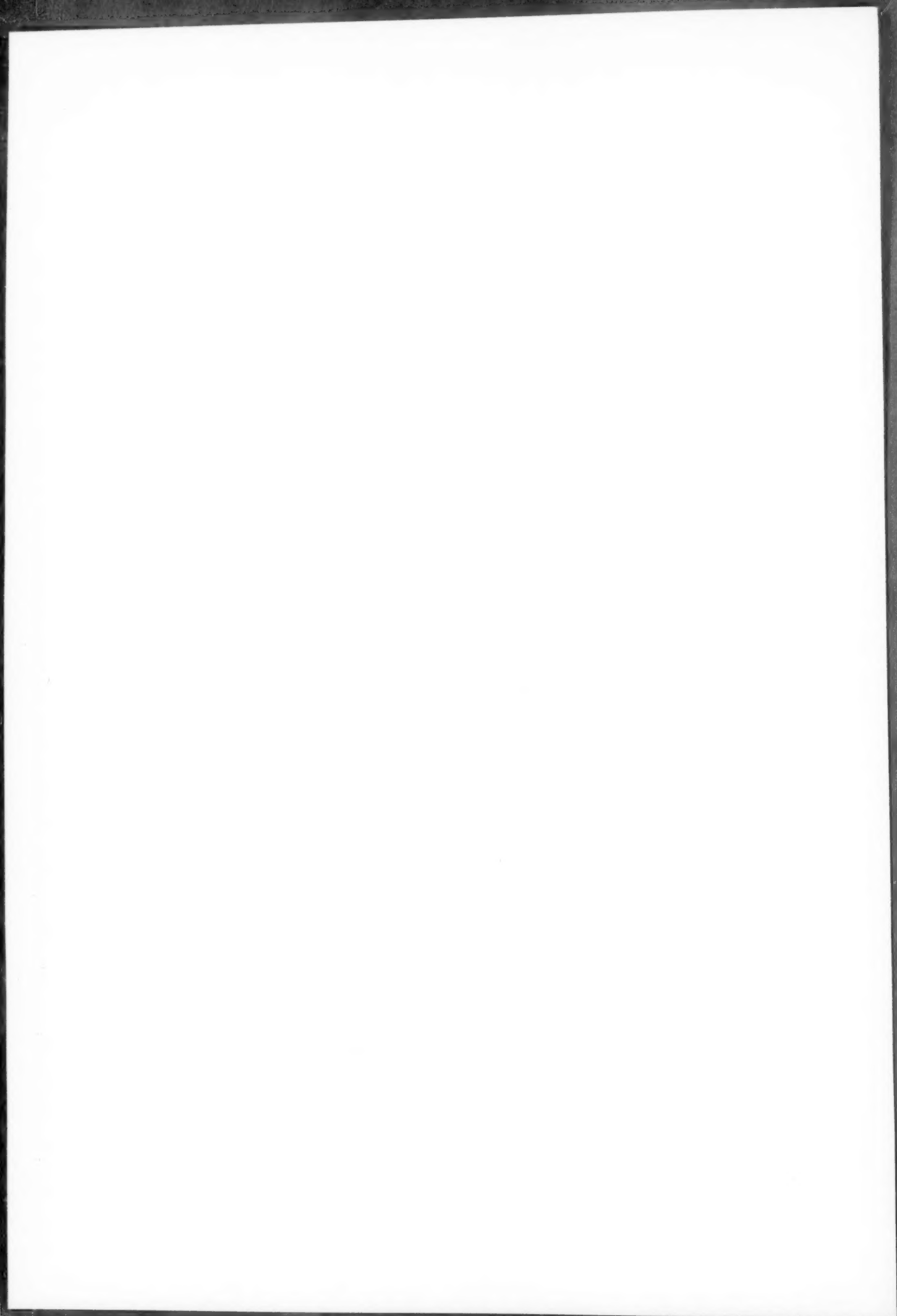
Covers the applications of psychology in the fields of business, industry, education, government, etc. Appears bi-monthly, February, April, June, August, October and December, the six numbers comprising a volume of about 480 pages.

Subscription: \$6.00. Single copies, \$1.25

Subscriptions, orders for single copies, and business communications should be sent to

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
NORTHWESTERN UNIVERSITY, EVANSTON, ILLINOIS





VOLUME 57
NUMBER 5

WHOLE No. 265
1944

Psychological review
Psychological Monographs

JOHN F. DASHIELL, *Editor*

**An Experimental Application of Projective
Principles to a Paper and Pencil
Personality Test**

By
Durham
HELEN SARGENT
Northwestern University

Published by
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

Publications Office
NORTHWESTERN UNIVERSITY, EVANSTON, ILLINOIS

1944

BF 1
.P8

COPYRIGHT, 1944

BY

HELEN DURHAM SARGENT

ACKNOWLEDGMENTS

THIS PROJECT has been made possible by an award from the James McKeen Cattell Grants-in-Aid Committee of the Psychological Corporation, and by a Northwestern University Fellowship for the year 1943-44.

Grateful acknowledgment is also due to Dr. Mary Phyllis Wittman, Supervising Psychologist for the Illinois Department of Public Welfare for arranging to have the "Test of Insight into Human Motives" administered at the Elgin State Hospital, and to Mrs. Ruth McCarn, Miss Lucille Scheuer, Mr. William Nimms, and Miss Katherine George of Northwestern University for cooperation in making follow-up data available for the student subjects.

The writer wishes also to thank Professors A. R. Gilliland, J. J. B. Morgan, R. H. Seashore, and W. L. Valentine of the Psychology Department at Northwestern University, for their volunteer services as judges, and to express appreciation to Professor Morgan for his direction of the research.

HELEN D. SARGENT

Northwestern University
Evanston, Illinois

vol. 57, no 5

TABLE OF CONTENTS

	<i>Page</i>
I. PROBLEM	
Introduction	1
Purpose	1
Basic assumptions	2
Fundamental questions	2
General approach	2
Guiding principles	3
II. DEVELOPMENT OF THE TEST	4
Title and general description	4
Construction of armatures	4
Make-up of the test	5
Technique of administration	5
III. SUBJECTS	6
Composition of groups	6
Attitudes toward the test	7
IV. PROCEDURES	9
Development of the scoring system	9
Feeling categories	9
Reliability Experiment I	10
Procedure	10
Results	10
Reliability Experiment II	10
Procedure	10
Results	11
Miscellaneous indicators	11
Conflict solutions	12
Method of scoring	13
Analysis of item neutrality	13
Procedure	13
Results	14
Total scores and subject profiles	14
Reliability Experiments III and IV	15
Self-agreement for one judge	15
Agreement between judges	15
Results	16
Inter-group comparisons	17
Alternate form comparisons	17
Single category comparisons	19
Combined score comparisons	20
Conflict solution comparisons	21
First person pronoun comparisons	22
Analysis of individual profiles	23

V. DISCUSSION OF RESULTS	26
Evaluation	26
Suggestions for further research	27
VI. SUMMARY AND CONCLUSIONS	29
Appendix A	
The test	31
Instruction sheet	31
Form 1 for men	31
Form 1 for women	32
Form 2 for men	33
Form 2 for women	34
Appendix B. Instructions to judges	37
Reliability Experiment I	37
Reliability Experiment II	38
Reliability Experiment IV	41
Appendix C. Examples of scoring and interpretation	43
Samples of answer scoring	43
Complete paper by subject EM4	43
Interpretive sketch	45
Appendix D. Replies of subjects to verbal questionnaire	47
Appendix E. Supplementary tables	51
Bibliography	57

AN EXPERIMENTAL APPLICATION OF PROJECTIVE PRINCIPLES TO A PAPER AND PENCIL PERSONALITY TEST

CHAPTER I

PROBLEM

INTRODUCTION

A CHALLENGING situation exists in the field of personality measurement. The so-called "paper and pencil" tests which include the various questionnaires, inventories, and rating scales purporting to "test" personality, have been criticized on a number of counts. The assumption that personality is a "sum-total" of traits which can be separately measured without reference to the whole personality has been attacked (1, 8, 14). It has been claimed that the statistical reliability of psychometric procedures is worthless because of failure to establish satisfactory validity in terms of predictive success (11, 12). On common sense grounds it has been argued that subjects can easily produce "good" or "poor" records at will. Furthermore, even the cooperative person has difficulty in fitting the subtleties of his own attitudes into rigid "Yes-No" answers. Although a high maladjustment score on such a test is usually significant (if for no other reason than that it marks an individual willing to admit characteristics which are unpopular according to group mores) a low score means very little. The subject may be well adjusted, adept at concealment, or his score may simply reflect his concept of what is "normal."¹

The projective approach to personality as a whole, functioning in more natural,

less circumscribed situations, instead of utilizing what have been termed "destructive-analytic procedures," (8) has stimulated new research.² The newer techniques are not, however, beyond criticism. Not only the problem of how to quantify and objectify, but the question of whether or not it should even be tried, has filled the literature with polemics (23). But in spite of experimental difficulties, many clinical psychologists prefer projective methods to other devices for personality study, for the reason that the former do apparently illuminate wishes, fears, and affective attitudes which are not readily verbalized even when the subject is fully aware of them. Practical experience has in many instances proved them to be a short and direct route to the collection of a great deal of biographical and personal material which would have required more time and skill to gather by the usual interview and observational methods (11, 12).

The obvious advantages of paper and pencil tests (using the term now in its literal sense) in convenience and low cost for group administration, and the strongly felt need for more penetrating diagnostic techniques in the "screening" of large groups, appear to justify an attempt to combine the two approaches.

PURPOSE

The purpose of this experiment has

¹In a recent reading of state hospital records, the writer noted that a large majority gained scores of "Excellent Adjustment" on a well-known inventory which has since been abandoned by the institution for this reason.

²For a comprehensive review of the historical background, theory, and applications of projective methods, see Sargent (23).

been to develop a personality test of the paper and pencil type which would embody the principles of projective methods. This rather ambitious program must, of course, be broken down into a series of steps, if a contribution to method is to result. The first necessity is a clear definition of the specific principles which are to be transferred from clinical experience to group application. These may be briefly outlined as follows:

BASIC ASSUMPTIONS

A projective test is based on the theory that (1) when a subject is confronted with a neutral, ambiguous stimulus (2) which requires a response, (3) he will react in terms of his own individual personality, and therefore it may be assumed that (4) whatever meanings he finds are his own, and hence are highly revealing of his personal characteristics and experience. Furthermore, since (5) the subject is unaware of the purpose of the purpose of the experiment, and cannot detect the factors on which he will be scored, results are less apt to be distorted by an effort to choose "right" answers. Finally, not only the content, but (6) the manner and form in which he perceives and organizes the material; his selective response to parts of the situation; and his voluntary and involuntary ways of expressing himself may all be regarded as potentially worthy of study.³

FUNDAMENTAL QUESTIONS

The second preparatory step is an outline of questions which must be answered if the feasibility of applying projective principles to a paper and pencil situation is to be determined:

³ The most commonly used projective stimuli have been inkblots, pictures, incomplete stories, and an assortment of materials from plastics to highly structured toys and representative objects (23).

1. Can a paper and pencil instrument be devised to meet the requirements of neutrality, meaninglessness and disguised purpose, which at the same time successfully calls forth a response from the subject?

2. Will the responses to such a "test"⁴ reveal individual differences between subjects?

3. Can a scoring system be evolved which is capable of identifying and objectifying these differences?

4. Can such a system be used reliably by different scorers and interpreters?

5. Can the information derived in this fashion be established as valid by such criteria as capacity to distinguish between groups of known composition, consistency with other personal data and results on other tests, and also by success in prediction?

6. Finally, can the interpretive material be handled in such a way that meaningful patterns appear which help to clarify the clinical picture of a single individual, in terms of both inter- and intra-individual comparisons?

Obviously the answer to these questions is the work of many researches, not of one alone. They involve both the collection of material for intensive case study, and the gathering of data from large, representative groups. For this reason, the present research has concerned itself chiefly with the first four questions, and has only touched upon the fifth and sixth.

GENERAL APPROACH

A somewhat different procedure has been devised than many which have been used in the development of other projective methods. It can, no doubt, be criticized as pedestrian. A frequent plan of attack has been to administer a new projective device to subjects about whom a great deal of information was at hand,

⁴ Quotation marks will not enclose the word "test" each time it is applied to the technique of this experiment, but these should be understood, since the term ought to be reserved for measures of proved reliability and validity.

and to interpret results in terms of such data (15). Although this approach is useful, it renders the methods themselves difficult to evaluate. Often interpretation has been based upon some established theoretical system, such as psychoanalysis (3), or a conceptualization of personality has been erected for experimental purposes (2). Checks on reliability of scoring and interpretation have been frequently ignored;⁵ claims of validity, based on congruence with other data have been sometimes hastily advanced.⁶ In the later stages of experimentation, blind matching and blind interpretation techniques have been used. Although these do succeed in demonstrating congruence between data, or between judges' interpretations, they offer no information on the process of interpretation or on the cues which have been used. The logical justification of this method, stripped of the holistic rationale which usually accompanies it, is that the possibilities of a method should be explored before a great deal of detailed work is put upon it. Hence the rough clinical method of evaluation is used first. Unfortunately, spectacular results are intoxicating and investigators are tempted to put a new device into immediate use, postponing the effort to fill in a foundation. It would be unfair to workers in the field to imply that such work is not now being done. It is underway (11, 12, 13, 18), but the fact that it has followed instead of preceding wide broadcasting of certain tests has invited disapproval and skepticism.

⁵ Murray and his coworkers tried several indices of reliability for judges' independent ratings, but a multiplicity of factors, including the use of the "conference method" in arriving at the assignment of "press" and "need" ratings, somewhat obscures their significance (20, P. 264, ff.).

⁶ See Macfarlane's criticism of this procedure (17) and Sargent's discussion of methodological problems (23).

The present investigation is primarily concerned with an effort to "pin down hunches" used in the interpretation of a certain type of projective product. It has proceeded on the following assumptions based on research which has gone before: it has been assumed, first that the mechanism of projection has been sufficiently demonstrated (19, 24) to warrant the expectation that it will operate in a test such as the one to be described; and second, that the exploratory research mentioned above has shown the value of similar techniques to an extent which makes possible an attempt to build a test on the same pattern using new materials. Such an extension of projective theory in no way seeks to supplant existing methods, but rather to test in a new medium the efficacy of the principles underlying them.

GUIDING PRINCIPLES

Finally, certain principles have been adopted. First, it was decided to make a strictly empirical approach to the data, freed as far as possible from theoretical biases of others, or of the investigator. Steps taken to ensure this will be described later in more detail. Second, it was decided to make the data themselves, in this case the written answers of the subjects, the point of departure, rather than the personalities of the writers. It was felt that if objectifiable differences could be noted consistently and reliably, the meaning of these in personality terms could be investigated later. Last, since established rules of test construction do not lend themselves to this type of problem without modification, statistical tools and methods of analysis have had to be adapted and forged as the project progressed. This has been a time consuming, but we hope not a wasteful effort.

CHAPTER II

DEVELOPMENT OF THE TEST

TITLE AND GENERAL DESCRIPTION

THE TITLE chosen was "Test of Insight into Human Motives." It is not, of course, a test of insight, either in the *Gestalt* sense of closure, or in the common meaning of understanding self or others. It was selected to arouse interest and to mislead subjects as to the purpose of the task set. The assumption is that most people pride themselves on understanding other people, and would hence be interested. Actually, the test consists of fifteen items describing persons in a conflict situation on which the subject is asked to write answers including "what did he (or she) do and why?" and "how did he (or she) feel?" It is presumed that people understand best those who are most similar to themselves, and that subjects are therefore prone to read into others their own motives and feelings. It is further assumed that the writer is more likely to "project" than to create, since it is an easier task to interpret from the standpoint of the one he knows best—himself.

The test items are called "armatures"⁷ because only the bare framework is presented without qualification or elaboration. Each describes an individual (of the same sex as the subject) in a problem situation with various courses of action left undefined. Instructions state that

⁷ The term "armature" has been borrowed from the fields of art and architecture, in which it denotes an adjustable framework on which the artist constructs his model. The dictionary definition of "armature" in its architectural meaning is: "any framework serving for stiffening or bracing." The more common word "skeleton" was considered for describing the bare bones of a situation presented in the items, but since a skeleton is rigidly set in a definite structure, and an armature can be adapted to the creator's purpose, the latter seems more appropriate for this usage.

"insight" into other people is essential for getting along with them and that the test is a measure of that capacity. It is also suggested that "insight" requires both imagination and ability to "put yourself in someone else's place."⁸

CONSTRUCTION OF THE ARMATURES

The armatures were developed as follows: A large number of situations were drawn up from everyday experience. In order to increase breadth of application, suggestions for many of them were adapted from Kimball Young's extensive chapters dealing with problems of personal adjustment (28). From the number devised, 36 were selected which most nearly met the following criteria:

1. The situation must present a conflict calling for some kind of action.
2. A minimum of emotion should be implied on the part of the characters, in order not to hamper the subject in reading into them his own.⁹
3. The characteristics of the imaginary persons should not be described, except as their actions may imply them.
4. As far as possible, the use of situations which suggest an answer in terms of popular stereotypes should be avoided.¹⁰
5. Proper names should not be used for the characters, in order to avoid possible affective associations with names which might be familiar.

⁸ See instruction sheet for test, Appendix A, P. 31.

⁹ A few exceptions to this rule were necessary; for example: in armature VIII on both men's and women's Form 1, parents are described as "disapproving" and "disappointed;" Armature IV of the same form describes a young couple as being "in love." In both of these, the emotions mentioned are an integral part of the conflict itself; in most instances the feelings of characters have been left undefined.

¹⁰ This was the most difficult criterion to meet, since the demand for wide applicability made also for triteness.

Of the 36 armatures which met these criteria, 24 were found adaptable for both men and women by a simple change in the sex of the leading character; six applied only to men, and six only to women. Accordingly, four forms composed of 15 questions each were made up, two forms for each sex.¹¹

In addition to selecting armatures which conformed to the criteria listed above, an attempt was made to ensure representation of six main areas of possible concern and conflict: family, opposite sex, social and friendship relations, vocation, religion or beliefs, and health.¹²

In the family area it was assumed that major conflicts would be in relation to dominance, obligation and affection. Armatures on this topic are intended as suitable mediums for the expression of reactions to such issues. Some deal with mother, some with father, others with parents. No specific items dealing with sibling rivalry are included although this might have been desirable. Armatures concerned with opposite sex relations involve conflicts concerning attraction versus judgment, separation, sex rivalry, and the like. In the friendship items, attitudes of clinging, submission to others, insecurity, dependence, or aggression are among those which have the most obvious opportunity to appear.¹³

¹¹ See Appendix A, Pp. 31-36 for copies of the forms.

¹² Table 12, Appendix A, p. 35, shows the number of armatures on each form which were classified in these areas.

¹³ Some of the armatures apply rather directly to war problems, for example, number XIV, F2W (in which a young man is about to leave on a "dangerous mission"), but except in one

MAKE-UP OF THE TEST

The test was presented to the subjects in mimeographed form with a separate sheet of instructions.¹⁴ In order not to place limitations upon length by demanding conformity to space requirements, blank sheets of paper were given out for the written answers.

TECHNIQUE OF ADMINISTRATION

Most of the subjects took the test in groups, although a few came in separately at odd hours, or entered the group after others had begun. An hour was allowed for writing, but this time limit was not strictly enforced. Those who were still writing when time was up, and who wished to continue, were allowed to finish. The group who took Form 1, after having taken Form 2 on a previous occasion, wrote for a slightly shorter period. Since we are more interested in the amount of material which a subject can and will produce in response to the stimuli, rather than amount in a stated interval, it would have been desirable not to use any time limit whatsoever.

When the tests were given out, whether to individuals or to groups, the instructions were summarized verbally and the subjects were encouraged to ask questions or express opinions either before or after the test. They were told that they might write on as few or as many items as they wished, in any order.

armature on F1M (XIII), war is not specifically brought in. It was thought that attitudes conditioned by war would inevitably be reflected in answers to other questions, and that it would not be advisable to limit even the trial forms by including problems which, it is to be hoped, are of only current import.

¹⁴ See Appendix A, Pp. 31-36.

CHAPTER III

SUBJECTS

THE "Insight" test has been given to 65 subjects, including 45 Northwestern University students (volunteers from beginning psychology classes) and 20 patients at the Elgin State Hospital. Of the Northwestern group, 19 subjects took both forms of the test, bringing the total number of papers to 84. Table 1 shows the distribution of subjects by group, sex, and forms of the test taken.

COMPOSITION OF GROUPS

In keeping with the objective of paying attention first to variations in the material itself, and with a view to working out a scoring system with as few preconceived notions as possible concerning the significance of the material produced, subjects were selected about whom nothing was known in advance, except that some were patients in a mental hospital, and others were students in a university. Abnormal subjects were included in order to ensure a sample in which extreme deviations might be anticipated. The two groups, which will be designated hereafter as

the NU and E groups are in no sense "controls" in the true sense of the term. Group comparisons which were subsequently made between the papers in the two groups were used as a rough measure of factors in the scoring which appear to hold most promise for further investigation. In interpreting the results it will be necessary to avoid generalization for either group as a whole, because of their assorted composition. It is conceivable that among the NU group there are prepsychotics or cases of severe neurosis. In the E group, on the other hand, certain cases about to be discharged may be considered relatively well adjusted.

The information presented in Tables 2 and 3 (Pp. 7-8) was gathered after the scoring, group comparisons, and interpretations had been completed. In the case of the E group (see Table 2) the investigator read the full hospital records. The data on the NU group (see Table 3) were collected with the assistance of University officials and counselors who furnished information from files and personal contacts. Personal inter-

TABLE I
Subjects classified by sex, group,* and form of test

	NU		E		Subject	Paper
	Men	Women	Men	Women	Totals	Totals
Form 1 only	6	8			14	33 F1
Both forms	5	14			19	
Form 2 only	4	8	13	7	32	51 F2
Sex totals	15	30	13	7	65	
Group totals	45		20		65	
Paper totals						84

* Northwestern University students (NU) and Elgin State Hospital patients (E).

TABLE 2
Elgin subjects: tabulation of data

Subject*	Diagnosis	Age	Education	Occupation	Intelligence#
EM ₁	D.P.U.	25	13	Student	106
EM ₂	S.M.E. w P.	38	9	Salesman	Good
EM ₃	D.P.U.	19	11½	Student	Average
EM ₄	C.A.	23	9	Truck driver	Good
EM ₅	P.P. no P.	16	7	Army	90
EM ₆	D.P.	29	11½	Factory worker	96
EM ₇	Par. state	32	8	Electrician	Poor
EM ₈	D.P.P.	43	14½	Unemployed	Very good
EM ₉	D.P.P.	29	16½	Draftsman	Very good
EM ₁₀	D.P.P.	25	12	Bus boy	Average
EM ₁₁	P.P. no P.	24	12	Unemployed	109
EM ₁₂	M.D.D.	31	9	Machinist	Good
EM ₁₃	D.P.U.	17	9	Messenger	78
EW ₁	P.P. no P.	17	8	Store clerk	70
EW ₂	D.P.U.	25	13	Stenographer	Very good
EW ₃	M.D.M.	15	11	Student	101
EW ₄	Obs. Comp.	37	16	Secretary	Very good
EW ₅	D.P.U.	36	8	Waitress	—
EW ₆	Con. Hys.	30	8	Nursemaid	Good
EW ₇	D.P.P.	26	14	Teacher	Very good

Summary

Diagnosis. Schizophrenics: paranoid (D.P.P.) 5, undetermined type (D.P.U.) 5; psychopaths without psychosis (P.P. no P.) 3; manic depressives: depressed (M.D.D.) 1, manic (M.D.M.), 1; psychoneurotics: conversion hysteria 1, obsessive compulsive 1; paranoid state 1, syphilitic meningeal encephalitis with psychosis (S.M.E.) 1; chronic alcoholic (C.A.) 1.

Age: Average for men 27 (range 17-43). Average for women 26.7 (range 15-37).

Education: Average for men 10.9. Average for women 12.3. (None had less than eighth grade; seven attended college.)

Intelligence: Superior, 5; "good" 5; average, 5; below average, 4, not rated, 1.

* Subjects were identified throughout by code numbers giving initial of group, sex, and subject number in group.

Figures refer to I.Q.'s on the Binet and Wechsler; word ratings are from the Elgin Initial Examination.

views were held by the investigator with 20 of the students. The latter group was also given the Guilford Personality Inventory and a Rorschach examination. Much of this material does not lend itself to tabular presentation, since both the amount and kind of information vary from subject to subject. The bulk of it is beyond the scope of this study, and is being reserved for future use.¹⁵

¹⁵ A study of inter-test consistency using the Guilford, Rorschach, and "Insight" test data is in progress with the assistance of Mr. Arthur Reinitz of Chicago State Hospital, who is making "blind" interpretations of the Rorschach protocols.

ATTITUDES TOWARD THE TEST

The personal interviews included a brief oral questionnaire concerning the subjects' reactions to the test situation. More than half of the group volunteered that they had used their own experiences and preoccupations in responding to the armatures; a number said they had used "experiences of myself and others," and only two claimed to have made an entirely objective approach.¹⁶ In response to the request to state which questions were most interesting, or best recalled,

¹⁶ This denial may, in itself, be significant, and is perhaps worth further investigation.

TABLE 3
Northwestern subjects: tabulation of data

College	Academic Grade						Totals
	F	S	J	S	G	Unc.*	
Liberal Arts	15	5	1	1			22
Education	2						2
Speech		1					1
Music				1	1		2
Journalism			1				1
Commerce	1						1
Summer School	2	2	1	1		7	13
Unclassified*	1					2	3
Totals	21	8	3	3	1	9	45

Summary

Age: Average for 30 subjects 18.4 (range 16-23); 2 students over 30; 13 unclassified.

Grades: Mean for 45 subjects 3.78 (range .75-6.5).**

Entrance: Mean M.A. scores for 29 subjects between B and C (5 A, 3 D). Unclassified, 29.#

* Subjects were asked to fill in age, class and school but many overlooked the request. These data are not on file at the University for summer students.

** Grades were figured in points (A 7, A- 6, B 5, B- 4, C 3, C- 2, D 1) for the quarter in which the test was taken.

The letters A to E represent quintiles on the Ohio State Psychological Examination.

12 spontaneously acknowledged that these were associated with personal matters, and several gave illustrations. Only two could remember none after about a week's time. The imaginative versus the moralistic approach was judged by a question which asked whether the subject had said what he thought the characters *would* do, or what they *should* do. "Would" was stressed by 60 percent, 35 percent said they had used "should," and 5 percent recalled a combination of the two.

Finally a general like-dislike reaction was solicited in a question phrased to make the subject feel free to express

either point of view: "Do you like to do this sort of thing, is it sort of a bore, or did you find it hard?" Fifteen liked the test with varying degrees of enthusiasm, ranging from "Oh, I just love to do tests like this—I like to get my ideas down," to a guarded "I liked it very much. I thought it was very interesting." Five expressed different degrees of distaste, complaining of vagueness, of "so little to go on," or of not understanding what the purpose of such a task could be.¹⁷

¹⁷ A verbatim report of the questionnaire replies, classified by questions will be found in Appendix D, pp. 47-50.

CHAPTER IV

PROCEDURES

THE HANDLING of the material actually involved seven separate treatments: four experiments concerned with the reliability of the scoring technique; a statistical analysis designed to check on the assumption of "neutrality" of the armatures; an analysis including statistical comparisons between groups on a number of factors; and a qualitative study of individual records in relation to supplementary data. The procedure and results of each study will be presented separately in the sections below which deal with the steps in developing the scoring, and its differential function. Discussion of the results as a whole will follow in the succeeding section.

DEVELOPMENT OF THE SCORING SYSTEM

The 14 papers for the first group of subjects who took Form 1 of the test were thoroughly studied and analysed in a number of different ways. Such factors as the number of questions answered, the number of lines written, the conflict areas emphasized, emotional words and phrases appearing more than once in a paper, the elaborations which the subject made upon the original theme of the armature, evaluative judgments, philosophizing and *clichés*, introduction of apparently irrelevant or senseless statements, types of conflict solution and frequency of alternate solutions, use of first person pronouns, and the appearance of material which could be scored by Murray's system of "press" and "need" were noted.¹⁸

¹⁸ The Murray system (20), although it added considerable insight, and seemed to result in a fairly adequate portrayal of the affective tone of

These various approaches met with different degrees of success from the standpoint of reasonable ease and consistency in application, and of apparent value for bringing out the characteristics of the papers. Many of the above categories were retained, and a few were dropped as the scoring evolved.

FEELING CATEGORIES

It was obvious that affective differences appeared in the papers. Certain subjects tended to use only a few phrases which related to feeling or emotion; others used many, and these appeared to fall into similar classes. Some tended to use a predominance of phrases referring to sadness, disappointment and various words of depressive coloring; others put down many references to avoiding trouble, such as "forget about it," "didn't let it bother him," and the like. On the basis of these observations, an attempt was made to work out a classificatory scheme using the least possible number of categories into which phrases from the 14 papers could be fitted.

As a result of considerable experimentation, a ten category classification of "feeling expressions" emerged. Here the term "feeling expression" is operationally defined as a word or phrase which a judge can identify as referring to some form of emotion.¹⁹ These expressions

a paper, was difficult to use because of the amount of overlapping in the categories and the large number to be applied. It was eventually abandoned and a simpler, though still faulty system of "feeling" categories was developed in its place for the present purpose.

¹⁹ No presumption whatever is made as to the nature of feeling or emotion, conscious or unconscious, physiological or psychological.

included references to feelings of the following kinds: frustration, aggression, submission, avoidance, depression, pleasure, negative attitudes, positive attitudes, anxiety and conflict, and rationalizations of feeling.²⁰

RELIABILITY EXPERIMENT I

1. Procedure

Four experienced psychologists were enlisted to act as judges. They were each given 408 phrases excerpted from their context in the papers and typed on 3 x 5 slips of paper. Typewritten instructions and a list of categories with subcategories illustrating type of feelings to be included, were furnished to each judge. After sorting, the judges were asked to enter their classification on the face of the slip.

2. Results

Contingency coefficients based on percent of agreement between each pair of judges were calculated from Vernon's tables (26).²¹ The coefficients ranged from $.46 \pm .051$ to $.812 \pm .017$ with a contingency coefficient for the average judge of $.68 \pm .032$.²² Except in the case of judge E who agreed least with the other judges, there was more agreement with the experimenter than with each other; a fact which might have been predicted from the former's greater familiarity with the classification. The highest co-

efficient obtained was between the experimenter and the psychologist with whom she had worked in close collaboration for a number of years. From these results, and from an examination of the categories assigned, it was apparent that the category scheme had not been defined with sufficient clarity to ensure the use of uniform concepts. Within the framework specified, each judge appeared to be using a fairly consistent classification of his own, but one which differed from those of the others. In spite of these discrepancies, there appeared to be an agreement which was substantial enough to warrant further experimentation.

RELIABILITY EXPERIMENT II

1. Procedure

In order to test the results of training judges in the use of the categories, a new set of four judges were engaged and paid for their time. All were beginning graduate students, two of whom had had clinical training. A new and more detailed explanation of the scoring scheme and a list of scoring categories were prepared and given to the judges in typewritten form to be studied in advance of the training period.²³ One three hour practice session was held, in which the following procedure was used. From the original set of 408 items, 100 had been selected, including 25 of the most difficult, 25 of those on which most agreement had been found among the first judges, and the remainder of medium difficulty. These were taken up in round table discussion and classified for practice. When the categories appeared to be well understood, each judge took home to classify a set of 200 additional items, also selected from the ori-

²⁰ For a full description of the types of feeling included in each category, see Appendix B, pp. 37-42.

²¹ The value of C for a given percent of agreement increases with the size of t , which is the number of elements to be matched. In this experiment $t = 10$ (number of categories) and $t' = 408$ (number of phrases). Vernon's table provides values of C and $S.E. \times \sqrt{N}$ for values of t from 1 to 12. C is practically equivalent to r when the grouping is finer than a 5×5 table (9, p. 391).

²² Table 27, Appendix E, p. 56, shows contingency coefficients between each pair of judges.

²³ See Appendix B, pp. 37-42.

ginal 408, and including items of assorted difficulty.

2. Results

On Reliability Experiment II, contingency coefficients between judges ranged from $.715 \pm .040$ to $.880 \pm .016$. For the average judge, C equalled $.810 \pm .024$.²⁴ These figures show a substantial increase over the results obtained in the matched judgments in Reliability Experiment I. A significant increase also appears in percentage of agreement, whether complete agreement, agreement of four out of five or agreement for the average judge in each experiment is taken as the measure.²⁵ The number of items on which there is no agreement was also reduced, although the difference is not statistically significant. In order to check on the possibility that the reduction in the number of items, rather than training and redefinition, accounted for the improvement, comparisons were also drawn between the same 200 items in both experiments. The percent of complete agreement is even more markedly increased by this procedure, although there is no significant decrease in the amount of total disagreement.²⁶

Further evidence for improvement in the second experiment was found by comparing the percentage of agreement in each separate category for both experiments.²⁷ Reliability Experiment II shows consistently better agreement in all categories except number 10 (rationalized feeling), which was used twice as frequently by the original judges. The latter

category appears to have been least clearly defined in the first experiment, and seems to have been used as a dumping ground for doubtful items.

On the basis of the first two reliability experiments, the "feeling expression" classification was considered sufficiently promising to retain. This was supplemented by several additional categories selected from those tried out on the original papers. These included seven miscellaneous indicators and ten types of conflict solution. After the statistical analysis of the papers, the former were revised and rearranged into two groups of categories: one consisting of three "cognitive" or non-feeling categories, and the other of three "maladjustment indicators" which appeared to distinguish the products of abnormal subjects (see Table 4).

MISCELLANEOUS INDICATORS

The following categories were temporarily named "indicators" as a matter of convenience, but no advance guess was made as to what their meaning might be:

a. Elaboration: elements added to the situation which are not stated or implied in the armature. For example, "The girl's father was himself a college graduate."

b. Evaluations: generalized comments of a valiative kind, such as "one should not live in the past," "one owes a duty to one's parents," or "it would be selfish for the boy to . . ." and the like.

c. Qualifications: statements suggesting reservations such as "it would depend," "that is, unless," or "except in case," as well as most phrases beginning "but if."

d. Subjectivism:²⁸ phrases or statements which appear meaningless in relation to the situation, the context or the feelings ex-

²⁴ Contingency coefficients and percentage of agreement between each pair of judges for Reliability Experiment II are presented in Appendix E, Table 18, p. 51.

²⁵ Table 19, Appendix E, p. 52, compares percentage of agreement in Reliability Experiments I and II.

²⁶ See Table 20, Appendix E, p. 52.

²⁷ See Table 21, Appendix E, p. 53.

²⁸ This category was broadly interpreted and included mildly inappropriate statements as well as neologistic phrases and "word salads."

pressed. Presumably these are not, in truth, meaningless to the subject, but only to the reader. For example, "Then he married the girl, as God's was the improper method."

e. Pronouns in the first person: all appearances of "I," "me," "mine," "myself."

f. Answers without content (AWC):²⁹ in-

chose to solve the situations were defined as far as possible in terms of action taken; that is, what the leading character was said to do. In general, these were conceived as varying kinds and degrees of action toward or away from the situation:

TABLE 4
Final list of categories

<i>Single scoring categories</i>	
"Normal" feeling	
1. Frustration	Cp. Compromise
3. Submission, dependence	Es. Escape, retreat
4. Avoidance, escape	Cf. Conflict, confusion
5. Depression	RF. Reaction formation
6. Pleasure, elation	Ex. Exploration
7. Negative attitudes	Un. Unreal solution
8. Positive attitudes	O. No solution
9. Anxiety and conflict	
10. Rationalized feeling	
Cognitive expressions	
El. Elaboration	
Ev. Evaluation	
Q. Qualification	
Indicators	
2. Irrelevant feeling*	
S. Subjectivism	
P. First person pronouns	
<i>Conflict solutions</i>	
At. Attack	
Re. Resistance	
Ac. Acceptance	
	<i>Combined scores</i>
	Totals per armature
	F. Total "normal" feeling score
	C. Total cognitive score
	I. Total indicator score
	Ranges
	FR. F range (lowest to highest per armature score)
	CR. C range (lowest to highest)
	IR. I range (lowest to highest)
	"Zero" scores
	OC. Number of "normal" feeling categories not scored
	OA. Number of armatures receiving no feeling score
	NA. Number of armatures answered
	F/C. Ratio of F total to C total

* Classified here according to apparent diagnostic significance on the basis of inter-group comparisons. (See Tables 6 and 7.)

volved or senseless answers which yield no other score.

g. Suspicion:²⁹ remarks of a suspicious nature whether projected to the characters or directed at the test itself, such as "he wondered how she spent her money" or "it is unfair to ask an answer on the basis of what is known."

CONFLICT SOLUTIONS

The various means which subjects

²⁹ Both suspicion and AWC were dropped for quantitative scoring, the first because of low frequency, the second because it could be better handled as an aspect of the total pattern. These were, however, qualitatively noted.

a. Attack: active entry into the situation, either to alter its unpleasantness, or to take full advantage of its pleasantness.

b. Resistance: active opposition to an encroaching situation, erection of barriers.

c. Acceptance: passive submission to difficulty or acceptance of benefits without activity on the part of the subject.

d. Compromise: discovery of ways to "get around" or modify the effect of a situation without actually altering it.

e. Escape: retreat of any kind, whether by rationalization or denial, or by physical withdrawal.

f. Conflict or confusion.³⁰ many courses

³⁰ It is important to distinguish these solutions

are considered but none are chosen.

g. Overcompensation:³¹ extreme response to the demands of the situation, overdoing compliance.

h. Reaction formation:³² adoption of an opposite extreme of action in response to frustration of an original drive.

i. Exploration: action chiefly of an investigatory nature; an effort to gather more facts, achieve more understanding, before acting.

j. Unreal solutions: introduction of an imaginative or unreal element which solves the situation without forcing the subject to meet it as stated. In the "Insight Test" answers, such solutions were always dependent upon an elaboration of the material.

METHOD OF SCORING

Every paper was typed in quadruple form. One copy was filed with the original; three were cut up for sorting by questions into equivalent sets. Each separate answer was identified by a code number which stood for form of the test, sex, group initial (NU or E), subject number, and number of armature. The purpose of this procedure was to avoid halo effect present in whole paper scoring; to obviate bias introduced by incidental information about the subject collected at the time of testing; and to permit comparison of all responses made to a single armature.³³

from answers in which two or more clear cut alternatives are stated, based on different premises. Presence of these were noted in summarizing.

³¹ Frequency of this solution was so low that it was dropped.

³² Brown (5, p. 173) describes the Freudian mechanism of reaction formation as a strengthening of repression by denial of conflict through behaviors opposed to the unconscious wish. It is closely allied to sublimation through which drive is diverted to socially acceptable channels. The former term was chosen for use here in order to avoid the evaluative connotation usually attached to sublimation.

³³ Child (6) has offered some helpful suggestions toward similar methods of increasing the objectivity of a single investigator dealing with ratings of qualitative materials. Although not directly applicable to this problem, some were adapted as guides in the procedure used.

The investigator adopted the following procedure for scoring the answers. First, each phrase which seemed to express feeling as already defined, was set off in brackets, and the number of the feeling category was entered above each. After feeling expressions had been identified, indicator scores were entered by a different set of symbols. Finally, the answer as a whole was read and parts directly describing action were underlined and identified by conflict solution category. The paper was then turned over and scores were summarized on the back.³⁴

ANALYSIS OF ITEM NEUTRALITY

1. Procedure

Before total scores for whole papers on each category could be computed, it was necessary to compare the stimulation power of the armatures themselves in order to determine whether these were actually "neutral;" that is, equivalent in producing the different scores. Since subjects had been allowed choice of items, a total score in a category might reflect this selection, rather than an individual tendency to project a particular feeling tone, or use an idiosyncratic approach. Since, if armatures are equally provocative, the present-absent frequency of a given category on one item should be in the same proportion as the present-absent frequencies for all taken together, the latter frequency provided a measure of expected proportionality. The differences between the observed and expected proportions in each category divided by their standard errors yielded a series of critical ratios for the 60 armatures on the four forms.

³⁴ See Appendix C, p. 43 for sample of an answer scored by this method.

2. Results

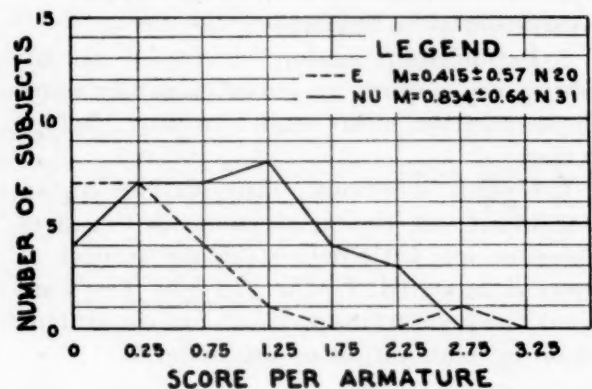
The majority of armatures on each form were found to differ significantly from expectancy based on the assumption of equivalence. All differences falling between the five percent levels of confidence (± 1.959 sigma) were attributed to chance and were arbitrarily assigned a score of 3. Differences which were significant between the five and one percent levels (± 2.575 sigma) were scored 2 or 4, according to the direction of the difference; the heavier weight being given to categories appearing on armatures which brought out significantly *less* of a particular feeling. Differences significant at the one percent levels were scored 1 or 5 according to direction.

The method is open to criticism on the grounds that some frequencies are small, and that actual sigma weights would be more accurate. It seems preferable, however, to a procedure which ignores the armature difference, and it appears sufficiently refined for data in which there is necessarily a large amount of error from other sources.³⁵

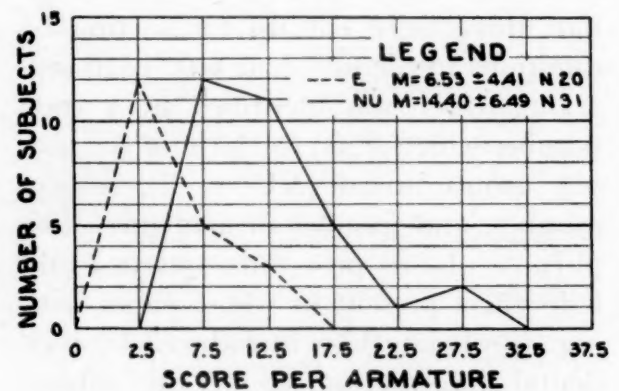
TOTAL SCORES AND SUBJECT PROFILES

Charts were drawn up for each subject showing the total weighted scores for each category.³⁶ The next problem was to derive for each subject a measure of comparative standing in the different scores. Before distributions could be made, it was necessary to allow for differences in the number of questions answered. Since the labor of making actual counts for each armature for all subjects would be prohibitive, a score per item was procured by dividing the

category total by the number of armatures on which the subject had written. Curves for these scores were then plotted for each category for the group as a



A. CATEGORY I. FRUSTRATION.



B. CATEGORY F. TOTAL FEELING.

FIG. 1. Typical curves showing distribution of feeling categories in E and NU groups.

whole and for each sub-group combination (E, NU, Form 1, Form 2, Men, Women). The curves differed from each other, but all shared the common characteristic of more or less marked positive skew. (See figure 1, showing representative curves.) For this reason, the means and standard deviations computed for use in inter-group comparisons were inappropriate for showing relative individual standing. The medians and quartiles were, therefore, computed as points of reference.

³⁵ See Appendix A, pp. 31-36 for tables of weights assigned to categories for each armature.

³⁶ See Appendix C, figure 7, p. 44, for sample chart.

On the basis of quartile standing, profiles were plotted on each chart, one for the single scoring categories, and one for certain combined scores, showing points of emphasis in a given paper. The latter scores were developed after inspection of the profiles indicated a rather consistent difference in pattern between the E and NU groups.

The combinations retained were as follows:

1. F total: a total feeling score per armature derived from the sum of weighted scores in categories 1 and 3-10 inclusive, divided by number of armatures answered by subject.

2. C total:³⁷ the total "elaboration" plus "evaluation" plus "qualification" score divided by number of armatures.

3. I total: the total score in maladjustment indicators "subjectivism" and "first person pronouns" plus score in feeling category 2, "irrelevant feeling" divided by number of armatures.

4. F range: the range between lowest and highest F scores as a measure of concentration or spread in type of affect.

5. C range: range between lowest and highest cognitive scores.

6. I range: range between lowest and highest maladjustment indicator scores.

7. Zero categories (OC): number of "normal" feeling categories in which no scores were given in an entire paper.

8. Zero armatures (OA): number of questions in which no expression of feeling occurred.

9. F/C quotient: the ratio between the F total and the C total.³⁸

The differential function of these various scoring approaches will be presented in the section on group comparisons. First, two further reliability experiments concerned with whole answer scoring will be presented.

³⁷ "C" stands for "cognitive" categories as opposed to "feeling" categories.

³⁸ In order to compute this ratio, scores were translated into standard scores by Hull's method of conversion into a common distribution with mean of 50 and sigma of 14 (9, p. 182).

RELIABILITY EXPERIMENTS III AND IV

The third reliability experiment tested the consistency of one judge (the experimenter) in repeated scoring by the consecutive and detached answer method; the fourth measured agreement between judge A (the experimenter) and two other judges in scoring whole answers. Although the procedures differed, results were treated by the same method and will be discussed jointly.

1. Self-agreement for One Judge

About three weeks after the scoring of all detached questions had been completed, ten papers were selected³⁹ for a repetition of scoring by the experimenter in the context of the whole paper.⁴⁰ The scorer used exactly the same procedure as before, except that tabulation was done in the margin beside each question and summarized at the end.

2. Agreement Between Judges

The services of two judges who had agreed most closely with the experimenter in Reliability Experiment II were enlisted to proceed with the scoring of complete answers. Another practice session was held, at which time the additional scoring categories were explained and the recording system presented.⁴¹ The two judges were each given

³⁹ Selection was made by the use of Fisher and Yates' table of random numbers, reprinted in Lindquist (16, pp. 262-264). To ensure a representative selection from both groups, it was decided to use five E and five NU male subjects. Numbers were read from the table and subjects who had answered four most frequently chosen armatures, whose numbers fell closest to the random numbers, were used.

⁴⁰ The memory factor in this task cannot be entirely eliminated, but considering the complexity of choices to be made and the fact that several hundred answers had been read, minimizes the weight attached to its influence.

⁴¹ One change in the feeling category list was introduced. Category 2, "aggression" overlapped with Conflict Solution 1, "attack;" and when expressed as feeling could be absorbed in

typed instructions,⁴² and sets of envelopes containing detached answers sorted by armature. Although the final comparisons used only four questions on ten papers, the judges were asked to score all answers to armatures I, III, IV, and V on Form 2 (chosen because the largest number of answers were available) in order to equate their sample experience with the experimenter's on the same armatures.

The chief problem in estimating either self-agreement or inter-judge agreement was in finding an adequate measure for comparison. Item for item agreement was somewhat better for the self-repeated scoring than between judges, but was not much above chance expectancy, due not only to variations in judgment but to the overlapping of categories and freedom allowed in number of categories to be assigned per answer. Agreement in the rejection of inappropriate categories, if counted, resulted in spuriously high percents of agreement, even when non-functional categories (categories used significantly less than chance by any judge) were eliminated. Yet to ignore this type of communal judgment defeated the major purpose of the comparison, which was not to compare single elements but whole patterns. For this reason, it seemed more fruitful to contrast the distortion introduced into individual profiles as a result of different judgments.

the categories of positive and negative attitudes. It was therefore dropped and the category "irrelevant feeling" substituted. Although this belongs with the "maladjustment indicators" and not with "normal" feeling expressions, numbering was not changed since judges had become used to the original numbers and confusion would have resulted. Table 4 (p. 12) provides a complete list of categories used for final scoring.

⁴² See Appendix B, pp. 37-42.

3. Results

Figures 2 and 3 (p. 18) show quartile profiles for all ten subjects scored by both methods by Judge A. Discrepancies between solid and dotted lines show differences between profiles based on detached and consecutive scoring. Agreement on quartile placement could also be shown by plotting the fourteen category scores on a four-by-four contingency table, but the profile method is preferable for showing location and direction of differences. Pattern distortions introduced by the scoring of three different judges A, F, and G is shown in figure 4 (p. 18) for two papers in which agreement is fair but not close. Table 5 (p. 20) gives contingency coefficients indicating the amount of agreement on each paper for the four sets of paired judgments.⁴³

Several implications are brought out by the profiles in figures 2, 3, and 4, and by table 5. Considering the highly qualitative nature of the scoring, and the many possibilities for divergent scores (see above, pp. 14-15) the agreement can be considered substantial. It is interesting to note that all of the judgments agreed better on the E than on the NU papers. The greater variety of expression in the normal group and the finer shades of feeling expressed made the latter more difficult to classify.⁴⁴ For the group as a whole, for the

⁴³ Different quartile norms were computed for interjudge comparisons based on distributions for the armatures used. (See Tables 26 and 27, pp. 55-56.)

⁴⁴ In the E group there is one paper, EM8, which all judges apparently found more difficult than the others. This was written by a paranoid whose devious means of expression made it possible to read into his answers almost anything, or nothing. Consequently agreement was less. In M3 comparisons it should be noted that Judge A agrees better with both judges than with self. Check on the scoring of this paper suggests that unconscious changes in concepts entered in when this paper was analysed which were not part of the investigator's original

two subgroups, and on most of the individual comparisons, self-agreement exceeded inter-judge agreement. Judge A agrees slightly better with Judge F than Judge G, but on the whole, agreement

INTER-GROUP COMPARISONS

1. *Alternate Form Comparisons*

Chi square tests show no significant differences in the distribution of category frequencies between groups who

TABLE 5
Reliability experiments III and IV:
Amount of agreement between quartile ratings
expressed in contingency coefficients*

Subject	Experiment III				Experiment IV			
	Judges A & A		Judges A & F		Judges A & G		Judges F & G	
	C	S.E.	C	S.E.	C	S.E.	C	S.E.
EM ₃	1.000	.016	.682	.121	.682	.121	.900	.051
EM ₅	.844	.071	.682	.121	.577	.156	.682	.121
EM ₆	.900	.057	.935	.037	.935	.037	.900	.057
EM ₈	.682	.121	.438	.177	.682	.121	.682	.121
EM ₉	.935	.037	1.000	.016	.973	.026	.973	.026
5 E	.895	.053	.830	.075	.820	.079	.830	.075
M ₁	.773	.092	.577	.156	.300	.196	.438	.177
M ₂	.844	.071	.682	.121	.577	.156	.577	.156
M ₃	.438	.177	.844	.071	.577	.156	.577	.156
M ₆	.844	.071	.844	.071	.577	.156	.773	.126
M ₇	1.000	.016	.577	.156	.900	.056	.773	.126
5 NU	.830	.075	.727	.097	.623	.097	.647	.109
All 10	.865	.064	.785	.089	.740	.089	.750	.102
<i>All papers</i>					<i>Exp. III</i>		<i>Exp. IV</i>	
Proportion of agreeing quartile matches					73%		63%	
Contingency coefficients					.855 ± .064		.762 ± .082	

* Corrected for broad grouping (9, p. 319).

is about the same between the three pairs. Inspection of the profiles in Figures 2, 3, and 4, illustrates the fact that the discrepancy is, in many instances, due to the relative amount of emphasis on a given factor, rather than to an actual contradiction in the direction of judges' estimates.

definitions, or of the instructions given to judges. Constant checks against such interference were found necessary by all three judges, necessitating occasional comparative rereading. Note that profiles based on four armatures (figure 4) are only slightly different from those based on whole papers (figures 2 and 3).

took the four forms, either when men alone, women alone, or both sexes are taken together.⁴⁵ The latter group did, however show a difference between Form 1 and Form 2 which approaches significance at the five percent level. There is, then, at least a suspicion that homogeneity of forms, or of groups who took them, is lacking. Contrast between the sexes becomes significant at the two per-

⁴⁵ Only normal subjects were used in comparing forms, since E subjects took only Form 2. For comparative frequencies and Chi square values, see Appendix E, table 22, p. 53.

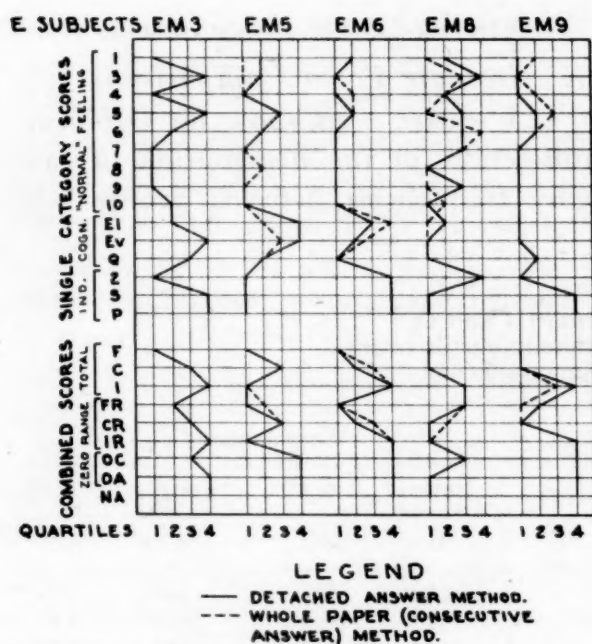


FIG. 2. Quartile profiles for five E subjects scored by two methods by one judge. (See tables 26 and 27 for norms.)

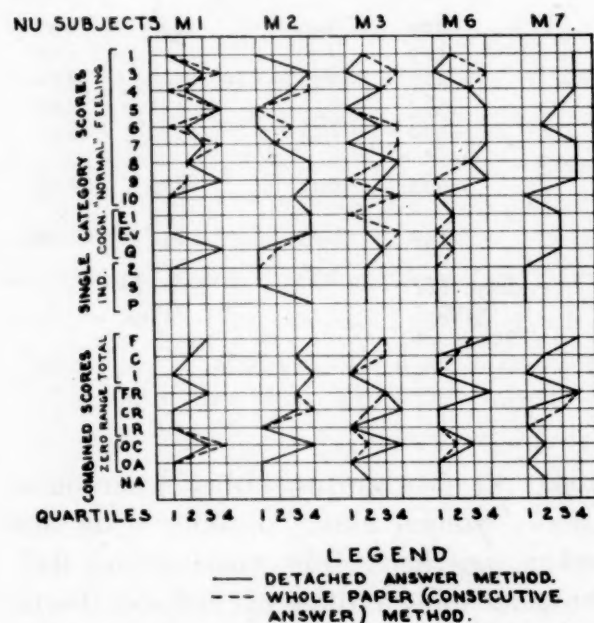


FIG. 3. Quartile profiles for five NU subjects scored by two methods by one judge. (See tables 26 and 27 for norms.)

cent level when both forms are combined in the comparison. Men on both forms show less depressive responses, and more avoidance reactions, the largest amount of difference appearing between the sex groups on Form 2. Either Form 2 was more effective in bringing out sex dif-

ferences, or the subjects who took this form were more dissimilar. It is conceivable that a larger number of cases might emphasize both the differences between forms and between sexes. For the present purpose, however, there appears to be too little evidence to warrant separate treatment in group comparisons.

Agreement between forms for the same individual is obviously a different matter than agreement for the group as a whole. When quartile profiles are compared for individuals who repeated the test, low agreement is expressed in low coefficients of contingency. For single scores, the C 's ranged from $.040 \pm .134$ to $.630 \pm .075$;

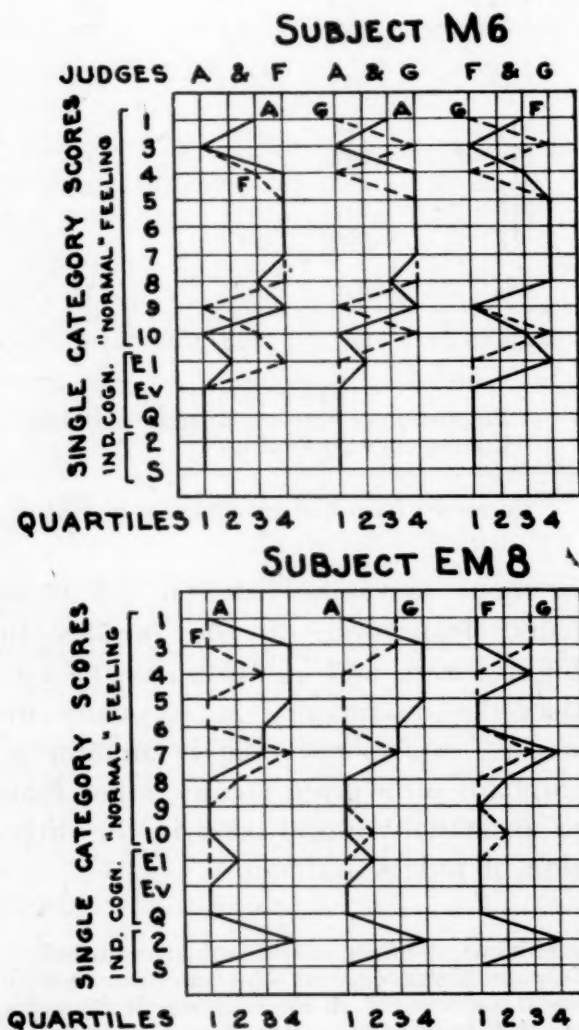


FIG. 4. Quartile profiles for one E and one NU subject on four armatures scored by three judges. (See table 26 for norms.)

for combined scores the range was from $.280 \pm .182$ to $.760 \pm .057$.⁴⁶ A similar finding results when F/C quotients are compared. The mean F/C quotients were 104.42 and 117.15 for Forms 1 and 2 respectively. Fisher's *t*-test results in a *t* of 1.7486 which, for 36 degrees of freedom, is not significant. When individual F/C scores on the two tests are correlated,

scores, *C* for the average subject was $.600 \pm .050$, and for single scores $.360 \pm .057$. This result is to be expected when categories are widened, since small individual differences are thereby obscured. Form to form differences do not invalidate a personality test, however, since a number of sessions may be required to bring out all significant aspects

TABLE 6
Differences in means of NU and E groups on single scoring categories

Category	NU Mean	E Mean	Dif.*	S.E. dif.	<i>t</i>	P>**
Feeling (F)						
1. Frustrated	.834	.415	.419	.176	2.380	.02
2. Irrelevant#	.082	.273	.191	.099	1.929	—
3. Submissive	1.007	.542	.465	.226	2.057	.05
4. Avoidant	.897	.401	.496	.175	2.834	.01
5. Depressive	1.347	.706	.641	.285	2.249	.05
6. Elated	1.565	.740	.825	.287	2.874	.01
7. Negative	1.764	.536	1.228	.338	3.633	.01
8. Positive	2.108	.722	1.386	.408	3.397	.01
9. Anxious	3.545	1.691	1.854	.438	4.232	.01
10. Rationalized	1.258	.488	.770	.402	1.915	—
Cognitive (C)						
El. Elaboration	1.660	2.140	.480	.466	—	—
Ev. Evaluation	1.983	1.419	.564	.592	—	—
Q. Qualification	1.649	1.283	.366	.250	1.464	—
Indicators						
P. First person	.221	2.011	1.790	.819	2.185	.05
S. Subjectivism	.012	.322	.310	.143	2.168	.05

* Italicized differences favor E group.

** Degrees of freedom 49.

Grouped with "maladjustment indicators" on subject profiles.

however, *r* is .260 showing an insignificant degree of relationship.⁴⁷

The contingency values presented above indicate that agreement was higher for combined scores than for single factors. These differences appear even more clear when the contingency coefficients corresponding to percentage of agreement between forms for the average subject on single and combined scores are compared. For combined

of personality, any one of which may be representative but incomplete.⁴⁸

2. Single Category Comparisons

Table 6 shows the results of Fisher's *t*-test applied to differences between category means for the E and NU groups.⁴⁹ No significant differences appeared be-

⁴⁶ Table 23, Appendix E, p. 54, shows contingency coefficients on quartile agreement between forms.

⁴⁷ See Lindquist's table 13 (16, p. 212) showing significance of *r* for small samples.

⁴⁸ See Tomkins' results on repetition of the Thematic Apperception test (25). Unfortunately it was not possible to have the E subjects repeat the test. Because of the greater reliability in scoring and certain consistent aspects of the abnormal pattern, it would be interesting to determine whether these scores are more or less stable than for normal subjects on a re-test.

⁴⁹ The *t* statistic tests the hypothesis that two

tween the means of the non-feeling or cognitive categories, but in all of the nine "normal" feeling categories, except "rationalization" which is evidently more closely related to the cognitive than the emotional approach, significant differences favor the NU group.

3. Combined Score Comparisons

Combined scores also show that the NU subjects produced a larger amount of

ceived no scores. As in the single score comparisons, scores depending on the cognitive categories did not differentiate the groups. The E group, moreover, used significantly more first person pronouns (see below, p. 22), more subjectivisms, and an amount of irrelevant feeling which just misses significant excess. The combined "I" scores which measure both amount and range of these indicators are also significantly greater. The most

TABLE 7
Differences in means of NU and E groups on combined scoring categories

Category	NU Mean	E Mean	Dif.*	S.E. dif.	<i>t</i>	P>**
F total	14.407	6.503	7.904	1.730	4.568	.01
C total	5.121	4.820	.801	1.380	—	—
I total	.251	1.921	1.370	.680	2.014	.05
F range	3.505	1.730	1.775	.687	2.583	.01
C range	2.070	2.190	.120	.299	—	—
I range	.039	.537	.498	.186	2.677	.02
Zero categories	.838	1.850	1.012	.377	2.682	.01
Zero armatures	.000	1.850	1.850	.460	4.021	.01
F/C quotient	1.070	.720	.350	.066	5.287	.01
Number of armatures	8.220	13.350	5.130	.097	5.288	.01

* Italicized differences favor E group.

** Degrees of freedom 49.

all varieties of feeling expression; that there was more variation and less concentration on one kind of affect alone (measured by the F range); that these subjects produced fewer answers in which no feeling appeared; and that there were fewer categories in which individuals re-

striking contrast between the group means is found in the higher average F/C quotient for the NU group. This difference is more than five times its standard error. The curve in figure 5 shows considerable overlapping, but it is

samples have been drawn from the same normal population, or from populations having the same mean and sigma. If significant at some pre-established level of probability, we are justified in rejecting the hypothesis of homogeneity. The *t*-test is better suited for small samples than the usual form of critical ratio or other large sample techniques, for the reason that *t* is differently distributed for each value of *n*, and hence does not go beyond the data but is "described exactly in terms of *t* itself, and does not require any estimate of a population parameter" (16, p. 54). Generalizations should not be extended to larger populations, since *t* simply expresses the probability of a given result for populations of similar size.

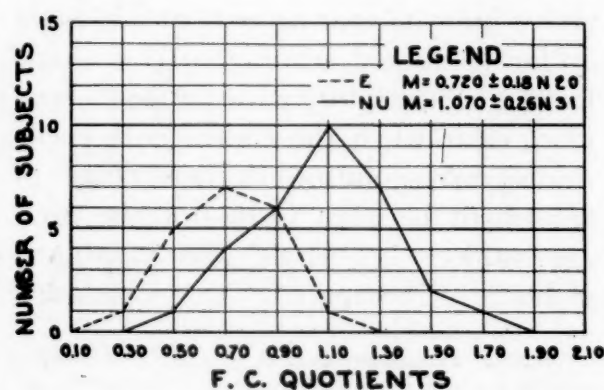


FIG. 5. Curve showing distribution of F/C quotients for E and NU subjects.

TABLE 8
Distribution of first person pronouns in groups and individual papers

Group N	Present	Absent	Number of subjects using			
			once	twice	three times	Excessively*
EM ₁₃	9	4	1	2	1	5
EW ₇	2	5	0	0	0	2
NUM ₁₅	6	9	3	1	1	1**
NUW ₃₀	7	23	4	2	0	1#

* Four or more times in the paper.

** Subject M₁₂. See Table 11.

Subject F₂. Follow up revealed frequent contacts with counselors in regard to adjustment problems.

interesting to note that although there were a number of low quotients among the NU group, only one in the E group rose above 100. A quotient of 106 was obtained by one psychoneurotic in the E group.

Since the E group tended to write on more of the armatures than the NU group (the difference is significant; see table 7, p. 20) a question is raised as to the influence of this factor upon observed differences. The greater number of items results, naturally, in a larger divisor for use in obtaining the score per armature, thus lowering scores for long papers. However, it seems reasonable to reject this explanation for several reasons. First, all comparisons are expressed in per armature feeling, rather than

total feeling. Second, the average length of the E answers per armature was equivalent to the length of the NU answers (see Table 24). Third, low feeling scores were found in E papers in which only a few items were answered, and conversely, high scores were found in some of the long papers. Finally, any constant factor affecting the feeling scores would have operated to the same extent on the cognitive scores; yet this did not occur.

4. Conflict Solution Comparisons

Since an all-or-none notation, rather than one of quantitative degrees, had to be used for the conflict solutions, the *t*-test is unsuitable for evaluating differences. The assumption that groups did not differ in the frequency of different

TABLE 9
Diagnostic classification of subjects using excessive first person pronouns

Subject	Number of first person pronouns	Diagnosis
EM ₂	17	Syphilitic meningeal encephalitis (without psychosis)
EM ₃	25	Dementia praecox (undetermined type)
EM ₄	91	Chronic alcoholism (without psychosis)
EM ₇	4	Paranoid state
EM ₉	4	Dementia praecox (paranoid)
EW ₁	33	Psychopathic personality
F ₂	5	Student*
M ₁₂	97	Student*

* Follow-up revealed that both subjects had presented personality and behavior problems. For fuller data on M₁₂, see table 11.

TABLE 10
Proportion of subjects in diagnostic groups using first person pronouns

Diagnosis	Number of subjects (Total N)	Number using first person
Dementia praecox	10	7
Paranoid	5	3
Undetermined	5	4
Psychopathic personality	3	2
Syphilitic meningeal encephalitis (no psychosis)	1	1
Paranoid state	1	1
Chronic alcoholic	1	1
Other classifications	4	0
Psychoneurotic	2	0
Manic depressive	2	0
University students	45	13

Number of cases too small for calculation of percentages.

conflict solutions (the null hypothesis) was tested by Chi square.⁵⁰ Chi square values which were significant at the one percent level justified the rejection of this hypothesis concerning the E and NU groups. Separate Chi square tests, equally significant, indicate that this difference is contributed by a preponderance of "Unreal" and "No solution" answers in the E group. A somewhat less significant difference (five percent level of confidence) also appeared between Form 1 and Form 2. Inspection indicates that this difference is attributable to the women of both groups. Here the significance arises from a number of small differences in various categories, rather than from emphasis on one or two specific types of conflict solution characteristic of either group.

5. First Person Pronoun Comparisons

Because of the interest of other investigators in this factor, and its obvious significance in these papers, tables 8, 9, and 10 (pp. 21-22) have been drawn up to show the frequency of this sign and its relative appearance in different diag-

nostic groups. Rosenzweig and Shakow found more self reference among schizophrenics than normals (22). Despert noted that the use of first person pronouns in stories was characteristic of the more disturbed children (7); Balken's anxiety and obsessive neurotics produced more than the hysterics (2); and Grings, using this criteria in "verbal summator" experimentation found that it distinguished between schizophrenic and depressive patients (10). In our E

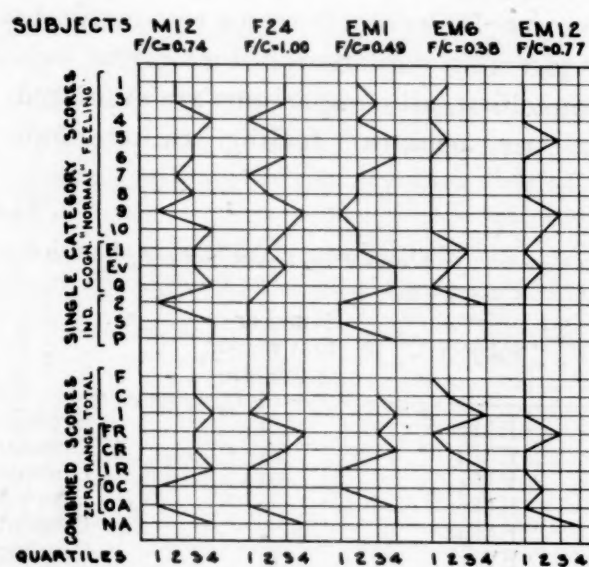


FIG. 6. Quartile profiles for three E and two NU subjects used for matching with supplementary data. (See tables 26 and 27 for norms.)

⁵⁰ See table 25, Appendix E, p. 55, for relative frequencies and Chi square values.

group, 70 percent of the schizophrenics produced this sign, including both the paranoids and the undetermined types. It was also present in two of three cases of psychopathic personality. It is especially interesting to note that only two of the normal subjects used an excessive number of personal pronouns (defined as more than four in an entire paper), and that the suspicion that these subjects were markedly maladjusted was later supported by data which counselors supplied (see note, table 9)

ANALYSIS OF INDIVIDUAL PROFILES

The meaning of the scores and their validity for individuals or for selected diagnostic groups is a matter for further research. Speculation is none the less tempting. It will, perhaps, be permissible to take a clinical rather than an

experimental view of the data at this point. On the basis of a quantitative and qualitative analysis of individual profiles, summaries and hypothetical sketches were written for all of the E subjects and for a number of the NU group.⁵¹ These cannot be presented here, but table 11 presents quotations from the original papers, includes comments from the analytical sketch written before anything was known of the subjects, and compares these with miscellaneous supplementary information for three E and two NU cases. Figure 6 (p. 22) shows the score profiles for the five subjects included in the table.

⁵¹ An example of the form in which these analyses were made, together with the score chart and original paper on which the sketch was based will be found in Appendix C, pp. 43-46.

TABLE 11

Qualitative comparisons between answers, analyses, and supplementary data for five subjects

Subject EM1, age 25, diagnosed paranoid schizophrenia, has been in hospital for two years and has not responded to electric shock treatment. Early life shows long history of family disharmony, hostility toward father and early obsessional tendencies. Has made a good institutional adjustment and has been placed on industrial therapy, doing routine clerical jobs at the hospital.

<i>Answers</i>	<i>Interpretive sketch</i>	<i>Hospital record</i>
Depending largely upon the personal affairs and abilities of the young man involved and the nature of the exigencies developed by his failure as well as his physical condition etc. one could say in taking the prerequisite conditions involved into consideration insofar as his future is concerned if it would suit his purpose he would comply. The reason for his complaisance is already implied.	Rationalizing tendency and low F/C quotient in subject with moderately high F level suggests possible paranoid traits, long, verbose, circular answers.	Diagnosed schizophrenia, paranoid type.
	Presence of some elaboration, moderate emphasis on 10 and excessive Q.	Rationalizes all past experiences. "Has an answer to everything."
	Lack of irrelevant feeling and subjectivism suggest S. is not deteriorated.	No deterioration.
	No solutions based on interjections of new or unreal elements.	Stream of thought always coherent.
His feelings would again depend largely upon the type of young man he was for it is not good policy to be presumptuous but we could normally assume that he should not be elated at failing.	Nowhere is there evidence of overt suspicion or blame projection, but the elaborate style, moralistic tone, inconclusiveness, caution, protectiveness, are paranoid in character.	Says he is illegally held, but there is no display of hostility.
	Vocabulary indicates good education.	Attended college. Student at time of breakdown.

TABLE II—Continued

<i>Answers</i>	<i>Interpretive sketch</i>	<i>Hospital record</i>
His feeling would depend on the strength of his attachment and the agreement of his actions and procedure concerning his mother	Stereotypy of expression. Repetition of phrase "it would depend largely" in reply to every question.	Fellow students thought him rigid.
Subject EM6, age 29, diagnosed dementia praecox undetermined type. Prognosis questionable, but shock therapy (electric) shortly after test was taken resulted in progress.		
It wasn't True and they had a child or children He felt "accused" for "something" he didn't know "anything" about.	Shape of curve indicates psychosis, probably schizophrenia. Predominance of unreal solutions suggests paranoid diagnosis, supported by emphasis on elaboration and qualification.	On admission showed marked push of energy, and persecutory ideas toward fellow employees. Secondary diagnosis, dementia praecox paranoid.
He felt normal . . . Must have felt normal . . . normal at times . . .	Striking perseveration of the word "normal."	Bizarreness, withdrawal from reality, delusional system not well organized.
He went to a "hospital" and was transferred to other hospitals instead of being released or told of his ailments. He felt normal but they transferred me to another hospital.	Autobiographical material.	Marked restlessness, excitement. Hydrotherapy ordered. Euphoria.
He must have been misused for the good can be the best according to the knowledge of a person, and still not be the best, and if he were bad he must have been taught the wrong way of living.	Guilt and its transformation into a paranoid mechanism are seen in statements about "misuse."	Homosexual seduction by employer at 15 caused much mental anguish. There is evidence of guilt feelings and depression.
He must have lost himself by the people through changes by the people who must have lost themselves by what they ate. . .	Paper gives strong evidence of deteriorating schizophrenia with paranoid components and guilt projection.	"I know my past. I am going to do something else, join the church and be a preacher. It's a transmitter that transmits high frequency voices, that's the way I get it."
He didn't do anything but kept going with her just the same. Until father committed suicide and later mother died then he lived with his brother and sister for a period of time until other situations came about.	Autobiographical?	Father was an alcoholic who shot his wife.
For he said a prayer to the lords, we use, and trod on those above. For he did read the Bible and told his wife it was worded wrong in places.	Religious delusions and pre-occupations are prominent.	Knows the Bible well as literature.
Subject EM12, age 31, diagnosed manic depressive depressed. Education one year of high school. Shock administered; discharged as recovered one month after test.		
<i>Answers</i>	<i>Interpretive sketch</i>	<i>Hospital record</i>
He dressed and changed his speech to be as much like the others as possible. When in Rome do as the Romans do.	Except for low F/C quotient and narrowing of F range, this might be a normal record.	Not hallucinated or deluded. Favorable prognosis.

TABLE 11—Continued

<i>Answers</i>	<i>Interpretive sketch</i>	<i>Hospital record</i>
He felt that his health was more important than his work He won't play second fiddle.	Low spirits and absence of compensatory phantasy suggest that this may be a depressive picture.	Lazy, no pep. Admitted because of acute agitated depression.
He feels cheap. It deflates his ego It makes him feel small and hurts his ego.	Several qualitative inadequacy feeling signs are present.	First depressive episode began with loss of "sex nerve."
Subject F24. The paper written by this freshman was selected from the women's NU group as most clearly indicating a neurotic state. Supplementary information quoted below was later obtained from her advisor.		
She tried to correct her speech and clothes so she would be accepted by her new friends.	Total feeling is somewhat below average and is concentrated in the frustration, dependent, depressed and especially the anxiety categories.	Very quiet.
She was naturally embarrassed . . . she also felt a little hurt.	Nearly every question contains references to fear, worry, shyness, etc.	She hardly ever speaks in class and when called on shakes like a leaf. Given a written assignment she writes copiously.
She would feel a little bad because she has disappointed her parents.	Many instances of moralistic, dutiful tone; high in evaluation.	On autobiography assignment wrote at length of closeness of family, fine home influence, valuable lessons learned.
She walked right out because she had an inferiority complex she felt funny and flushed.	What caused inferiority feeling?	Said all the important steps in her life had come as the result of suggestions by members of her family. Thinks it best for families to be close.
She feels that she would rather stay home with her parents and those people she knows well than go out with boys she does not know quite so well and is afraid she will not be liked by her dates for fear of giving them the wrong impression of her.	Probably self-conscious with the opposite sex. Seems impossible that a "glamor girl" could have written this paper!	Informant comments that most of her decisions seem to have been made for her. Wider environment of college is probably difficult.
Subject M12. Of all the NU subjects this freshman man's profile pattern (see figure 6, p. 62) most resembles those which predominate among the E papers. Through university sources the supplementary data was supplied.		
If I were in this situation I would forget about it. He should not go but if he likes the friend better who has the date he'll go making some lame excuse . . . with a twinge of remorse if he has not done it often before. After all it's his own life.	Excessive use of first person pronouns. Combined score profiles resemble the reversed "s" shape which characterizes the E group All C scores are high: evaluation, qualification, and elaboration. F/C quotient low for NU.	Described as "almost a pathological liar." Caused much annoyance to dormitory mates, although he was very well liked and others tried to help him. Inclined to brag about his past exploits.
If I were in this situation I would simply not care to discuss the matter.	Feeling emphasis is on avoidance and rationalization.	
I would at another time, if she offered me no chance, ask what the reason was. If she offered no good examples, I'd try to date her again. If she acts coldly I'd look for another girl.	Vocabulary good. Much of content is smoothly written and reasonable. It is the pattern itself rather than content that suggests a problem.	Excluded twice for low marks. Entrance score D, but opinion expressed is that he does not lack ability and intelligence.

CHAPTER V

DISCUSSION OF RESULTS

RESULTS of a number of procedures, each one leading on to a subsequent step, have already been presented in the preceding sections. It is necessary how to evaluate them in relation to the questions posed at the beginning of this experiment (see above, p. 2).

EVALUATION

The first question related to the problem of constructing a paper and pencil test which met certain projective requirements, including neutrality of subject matter; that is, sufficient ambiguity of the material itself to justify interpretation of response variations in terms of a subject's own attitudes, rather than as resulting from characteristics of the stimuli. Analysis of the armature items on the "Insight Test" indicated that complete neutrality could not be assumed. Certain categories appeared more frequently in response to certain questions than would be predicted on grounds that only chance factors were operating. Although a test of the verbal type is more vulnerable to the influence of meaning factors than pictures, ink-blot and other less structured materials, the finding calls into question an assumption of equivalent stimulus-value which has been given scant attention except in the Rorschach test and the TAT.⁵²

The second and third questions were concerned with the possibility of developing a scoring system to deal with individual differences, should they appear.

⁵² Beck's forthcoming book (4) presents detailed card by card data concerned with forms and details frequently encountered in the responses. Murray, in the TAT manual (21) gives "need" averages for each variable, but Harrison suggests frequency tables of common themes for each picture (12).

Although the system used is complex and in need of revision, it was successful in objectifying certain characteristics which varied from subject to subject and which could be treated in a quantitative fashion. It is particularly worth noting that the scoring factors are more useful in combination than separately, and that an analysis of the patterning of scores for individuals and groups yields more promising results than evaluation of single factors. An examination of the graphs in Figure 2 and 3 (see above, p. 18) reveals a general difference in the combined score profiles for "normal" and "abnormal" subjects. The combined score curve for the NU subjects in most instances roughly approximates an "S" in shape; for the E subjects the "S" is reversed.

The four reliability experiments demonstrated an encouraging amount of agreement between judgments both on the classification of "feeling phrases" taken out of context, and on the categorization of whole answers. The results of the first and second reliability experiments also showed that agreement could be increased by training judges in the use of the concepts. As might be expected, when judges were asked to manipulate a larger number of categories, rejecting and choosing as many as they saw fit, disagreement increased. On the other hand, better agreement appears when patterned scores are compared than can be shown on item for item comparison. Improved reliability, or at least applicability of more accurate means for determining it, might result through certain changes in scoring technique. The system itself might be sim-

plified; scorers might be limited as to the number of categories assigned per answer; or degree ratings on each of several categories might be required. A crucial question in regard to the last two suggestions is whether such an increase in statistical precision might not be achieved at the sacrifice of interpretive freedom and completeness. A solution to this problem might lie in the development of a simple set of quantitative scores for group screening purposes, supplemented by a qualitative guide for use when more complete individual analysis would be desirable.

The fourth and fifth questions concerned validity, a problem for which this experiment was not designed. The probability of overlapping between the "normal" and "abnormal" groups used, and their uncontrolled composition, make generalizations for either unsound. The group comparisons did, however, bring out differences which were statistically significant for the small samples used, thus indicating factors on the test which seem worth further investigation. Little difference appeared between the groups either as to the variety or kind of feeling expression, but a striking difference of degree appeared both for the feeling categories separately compared, and for all in combination.

The low affect characteristic in a psychotic group might easily be anticipated, if the entire abnormal population had been schizophrenic. It is interesting to note, however, that of five E subjects whose F totals fell in the upper quartiles, only two were diagnosed schizophrenic. Diagnoses for the other three were manic, obsessive-compulsive, and paranoid state. In the low affect group are to be found, in addition to schizophrenics, three psychopaths, one manic depressive depressed, a paretic, an alcoholic, and a

conversion hysteric. Analysis of intelligence and educational data is equally inconclusive. One of the high affect subjects was a college student at the time of commitment, and two were secretaries; on the other hand, the fourth was a low grade individual of secondary school education, and another was a high-school student of average mentality. Maladjustment factors producing affective repression in the test situation, not common to any one particular disease entity, seem the most likely tentative explanation.

Data on the sixth question, which concerned the possibility of individual analyses by comparison of the test products with supplementary information, have been presented here chiefly as a matter of interest and suggestion. The results are obviously inconclusive, but serve to indicate possible approaches to the interpretation of individual patterns. In this connection, the discrepancy between forms on individual retests should serve as a reminder that one small sample is insufficient as a basis for personality analysis, unless it can be shown by further experimentation that differential signs remain stable. Only normal subjects repeated the test in the present experiment, and among these were none whose patterns showed marked deviations. A not impossible outcome might be that close reproduction of a pattern on a second test might itself be a pathogenic sign. This is, of course, sheer speculation.

SUGGESTIONS FOR FURTHER RESEARCH

Several improvements on the technique might be developed from these findings. The name of the test and its general make-up appear to have been satisfactory for its intended purpose, judging from results cited and the response of the subjects. The number of

questions could conveniently be reduced to not more than ten. Whether to eliminate non-neutral armatures or to continue allowance for them by weighting is debatable. Interest is increased by meaning. The important necessity is not so much perfect equivalence in basic provocative power, as information in regard to stimulus differences for use in the treatment of results, and for dictating the elimination of items which prove to be insufficiently flexible.

The discarding of separate feeling categories for quantitative scoring is probably justified both on grounds of efficiency for group use (which involves the laborious scoring of many papers) and because their chief significance is in the portrayal of individual patterns rather than in the differentiation of groups. If they are retained for the latter purpose, certain revisions would be necessary. The "avoidance" and "submission" categories are constantly interchanged and confused in practice and might be subsumed under a "withdrawal-dependent" category. "Frustration" and "depression" might well be combined, and the "pleasure, elation" category included with "positive attitudes." "Rationalization" should be counted in the evaluation score, since it appears more closely related to the cognitive than to the affective categories. "Irrelevant feeling" and "subjectivism" are closely related, and should probably form a single classification.

For screening purposes, a simple count of feeling phrases, in relation to the number of elaborative, evaluative, and qualificatory expressions, plus the easily identified and highly significant first person

pronoun might prove practical. These scores, combined with some measure of subjectivism as suggested above, appear to provide a promising approach to differential scoring on the basis of the significant differences in F scores, F/C quotient, and maladjustment indicator totals found in the statistical analysis of these papers.

The scoring used for this experiment by no means exhausts the possibilities for quantitative evaluation. A comparison of this scheme with results obtained by Murray's "press-need" system (20), given a more thorough trial than it has been granted here (see above, p. 9), would be worth while. The formal criteria which Balken has used for Thematic Apperception protocols would also be well suited to these productions (2).

An effort to validate the "Insight Test" in its present form, with scoring unrevised, is not indicated by these results. If, however, a follow-up study of significant factors bears out the diagnostic usefulness of revised scoring, a series of validation experiments might profitably be undertaken. Comparisons with other projective test findings, with results on traditional paper and pencil inventories, and with full clinical data, should be included in the program. Prognostic predictions, both for hospital patients and for doubtful normals, whose subsequent progress could be followed, might be made and checked against later adjustment. This work is for the future. The present exploratory study has done little more than to stake out the site and spade up the topsoil of an area in which profitable research might be founded.

CHAPTER VI

SUMMARY AND CONCLUSIONS

THE POSSIBILITIES for developing a projective paper and pencil test for group and individual use have been explored. A list of conflict situations was presented to 45 college students and to 20 patients in a mental hospital as a "Test of Insight into Human Motives." The subjects were asked to write in any way they wished on the subject, "What did he (or she) do and why?" and "How did he (or she) feel?" Four forms of the test, two for men and two for women, consisting of 15 armatures each, were used.

A scoring system consisting of ten "feeling categories," seven miscellaneous "indicators," and ten types of conflict solution was developed on an empirical basis. After comparison of these factors had been made between the groups, the scheme was revised to include nine "normal" feeling categories, three cognitive categories, three maladjustment indicators, eight combined scores, and a quotient showing relationship between the total feeling and total cognitive scores.

The reliability of the scoring technique was subjected to four experiments, the first concerned with agreement on feeling categories only, using two different sets of judges, one set "expert" but untrained in the specific concepts, the other less experienced but given special instruction. The third experiment dealt with the self-agreement of one judge using different methods of whole answer scoring, and the fourth tested agreement between three judges scoring all replies to four of the Form 2 armatures. Before total scores were developed, an effort to equate the stimulus value of the armatures themselves in producing the scores

was made by a system of weighting.

Differences between means of the hospital and college groups were subjected to Fisher's *t*-test on 24 scoring factors. Charts were drawn up for individual subjects, and tentative analyses of individual patterns were attempted. The latter were checked against data gathered for all of the institutional cases and for a few of the University group.

The main conclusions to be drawn from the findings in the experiment are as follows:

1. Since more than half of the questions varied in their tendency to produce different types of affective expression, the neutrality of items, such as verbal armatures, cannot be assumed without check. It has been pointed out in the discussion that this finding has possible implications for other projective techniques.

2. Inconsistencies in the patterns of normal subjects who repeated the test emphasize the fact that one small sample of thought content is insufficient for personality analysis. Further study is needed to determine which, if any factors remain stable on a retest and what the significance of these may be. The four forms of the test used in this project differed little from each other as far as group frequency distributions for categories were concerned, but wide variations were found on individual retests of 19 university students.

3. The reliability of the scoring system, measured by agreement between judges in applying it, was reasonably satisfactory, in consideration of the complexity of the system and the qualitative nature of the data. The possibility of improved agreement through training of

judges was demonstrated.

4. Defects in the present technique regarding length of test, overlap in the scoring categories, and the inclusion of non-differentiating factors became evident. On the basis of present experience, simplification and greater effectiveness appear feasible to achieve.

5. Certain significant differences were found between papers written by mental patients and college students. Although interpretation of their import is guesswork, because of the uncontrolled method of subject selection, these differences do indicate certain scoring factors which deserve further attention. Most decisive of the group differences were the following:

a. Patients, as a group included a proportionally smaller number of feeling expressions of all kinds except "irrelevant feeling" on their papers. As far as the data go, there is no evidence to support the hypothesis that this characteristic is associated with any one diagnostic classification, or with education and intelligence.

b. The groups did not differ in the amount of elaborative, evaluative or qualificatory expressions, but they did differ in the ratio between feeling and cognitive scores. The mean F/C quotient for the mental patients was markedly lower than for the students. Although there is overlapping between groups, the difference is five times its standard error.

c. Several "maladjustment indicators" distinguished the hospital answers, including the appearance of "irrelevant feeling," a type of statement which was named "subjectivism," and excessive use of first person pronouns. The latter becomes especially interesting and seem-

ingly important when its appearance is broken down into frequencies in which it occurred in the papers of single subjects belonging to different diagnostic groups. For the subjects of this experiment, the appearance of this sign more than four times in a paper was without exception associated with maladjustment as judged by corroborative evidence.

6. On the whole, the results offer strong evidence that the mechanism of projection operates in a paper and pencil situation of the type used. This conclusion is supported by individual differences appearing in the patterning of scores, by the differential value of certain scores described above, and by the results of matching material from the papers, and from blind comments based on the answers, with data from supplementary sources. The latter demonstrates, on a qualitative basis, at least the probability that personal attitudes and experience are reflected in a way which makes them accessible for study and interpretation.

7. The exploratory nature of the experimental design, and the tentative character of the objective scoring used, demand qualification of conclusions, and caution in the extension of any generalization beyond the data of this project. On the other hand, the "Insight Test" seems sufficiently promising to warrant additional experimentation. Scoring possibilities have not been exhausted, and validation procedures might profitably be undertaken. The actual value of the test remains to be determined. As one writer has aptly remarked: "It is not unduly difficult to make up a personality test, but it is a gigantic undertaking to find out what it is worth" (27, pp. 228).

APPENDIX A

THE TEST

This section contains the test itself, consisting of the instruction sheet and copies of the four forms; table 12 showing the distribution of the armatures by area of conflict; and tables 13-14 giving the category weights assigned in scoring responses to different items. On the test copies, it will be noted that two sets of armature numbers are used. Arabic figures correspond to order on the mimeographed forms; Roman numerals in parenthesis are serial numbers which are the same for corresponding items on the men's and women's forms. Numbers I-XII on each form have counterparts on the form for the opposite sex.⁵³ Numbers XIII-XV are adaptable for one sex only. The serial numbers were those used in all handling of the data, and it is these to which references are made in the tables, charts, and discussion.

TEST OF INSIGHT INTO HUMAN MOTIVES

Instructions

Insight into other people helps us to get along with them. This is a test of your ability to "see into" others. This requires both imagination and the ability to "put yourself in someone else's place."

On the following pages you will find a number of situations described very briefly. After each you will find two questions which you are asked to answer in writing on the blank pages provided. Please number your answers to correspond to the questions on which you are writing.

Notice that the persons involved in the situations are not described. This allows you to use your imagination as to what sort of people the characters might be.

There are no right and wrong answers, but your explanations should show understanding of the characters as you see them. You may write for an hour. *Write first on*

⁵³ Through a clerical error the system breaks down at one point on the present forms. No. IX on F1M is the same as No. II on F2W; likewise No. IX on F1W is equivalent to No. II on F2M. The interchange was not discovered until after administration, but is more confusing than disastrous, inasmuch as other revisions are necessary before further use.

the items that interest you most, since you may not have time to answer all the questions in the time allowed.

FORM 1—Men

- (I) 1. A young man is working or studying away from home gets a letter from his mother, after the death of his father, asking him to move back home.
 - a. What did he do and why?
 - b. How did he feel?
- (II) 2. A young man has acquired religious and political opinions away from home which are in direct conflict with his parents' ideas. He is home for a visit and religious and political subjects are discussed.
 - a. What did he do and why?
 - b. How did he feel?
- (III) 3. A young man falls in love. In order to marry he must give up his studies and make some money immediately.
 - a. What did he do and why?
 - b. How did he feel?
- (IV) 4. A young man gets a good deal of razzing because he spends his week-ends at home instead of dating.
 - a. What did he do and why?
 - b. How did he feel?
- (V) 5. A young man's father has always looked forward to having his son take over his business and has educated him for it. The son becomes interested in another vocation.
 - a. What did he do and why?
 - b. How did he feel?
- (VI) 6. A young man discovers that a girl to whom he is engaged has had a very bad reputation in the past.
 - a. What did he do and why?
 - b. How did he feel?
- (VII) 7. A young man is caught in petty theft. His employer offers to let him keep his job if his parents will vouch for him. Otherwise he will be forced to look for another job without references.
 - a. What did he do and why?
 - b. How did he feel?

- (VIII) 8. A young man has a date with a girl he has been seeing regularly. At the last minute she breaks it without explanation.
a. What did he do and why?
b. How did he feel?
- (IX) 9. A young man whose parents are divorced is asked to testify against one in a law suit involving his own inheritance.
a. What did he do and why?
b. How did he feel?
- (X) 10. A young man has promised a friend to help him with some work. Another friend asks him to go on a blind date with a girl who is said to be very attractive.
a. What did he do and why?
b. How did he feel?
- (XIII) 11. A young man is offered a deferment in order to finish professional training. All of his friends are in the army or navy.
a. What did he do and why?
b. How did he feel?
- (XIV) 12. A young man has been captain of his high school football team, but fails to make the squad at the University. He is offered a scholarship at a small college.
a. What did he do and why?
b. How did he feel?
- (XV) 13. A young man is strongly attracted to a girl who is "above" him in wealth and social position. He hears indirectly that she is interested in him.
a. What did he do and why?
b. How did he feel?
- (XI) 14. A young man gets the impression that others are discussing him. On several occasions he thinks conversation has stopped or the subject changed when he entered the room.
a. What did he do and why?
b. How did he feel?
- (XII) 15. A young man has been feeling quite ill but the doctor tells him he is not seriously sick and it will do him no harm to go on leading a normal life.
a. What did he do and why?
b. How did he feel?
- father's death, asking her to move back home.
a. What did she do and why?
b. How did she feel?
- (II) 2. A girl has acquired religious and political opinions away from home which are in direct conflict with her parents' ideas. She is at home for a visit, and religious and political subjects are discussed.
a. What did she do and why?
b. How did she feel?
- (XI) 3. A girl gets the impression that others are discussing her. On several occasions she thinks conversation has stopped or the subject changed when she enters the room.
a. What did she do and why?
b. How did she feel?
- (IV) 4. A girl is disapproved by her friends because she spends her week-ends at home instead of dating.
a. What did she do and why?
b. How did she feel?
- (V) 5. A girl's parents have always looked forward to having her follow a particular career and have educated her for it. She becomes interested in something else.
a. What did she do and why?
b. How did she feel?
- (VI) 6. A girl discovers that a man to whom she is engaged has had a very bad reputation in the past.
a. What did she do and why?
b. How did she feel?
- (VII) 7. A girl is caught in dishonesty. Her employer offers to let her keep her job if her parents will vouch for her. Otherwise she will have to find other work without references.
a. What did she do and why?
b. How did she feel?
- (VIII) 8. A girl has a date with a man she has been seeing regularly. At the last minute he breaks it with no explanation.
a. What did she do and why?
b. How did she feel?
- (IX) 9. A girl who has been very popular in her home town is teased by new friends about her clothes and speech.
a. What did she do and why?
b. How did she feel?

FORM 1—Women

- (I) 1. A girl who is working or studying away from home gets a letter from her mother, after her

- (X) 10. A girl has promised a friend to help her with some work on a certain evening. She is later asked by another friend to go on a blind date with a man who is supposed to be very attractive.
a. What did she do and why?
b. How did she feel?
- (XIV) 11. A girl who has been studying away from home signs up for a summer job. When she goes home for a week-end she finds that her father had planned to take her on a trip with him.
a. What did she do and why?
b. How did she feel?
- (III) 12. A girl has an excellent chance for advancement in her profession. She becomes engaged and her fiance asks her to give it up.
a. What did she do and why?
b. How did she feel?
- (XIII) 13. A girl has planned to go away to a small college and to room with her best friend from home. Toward the end of the summer she is offered a scholarship at a large university.
a. What did she do and why?
b. How did she feel?
- (XV) 14. A girl's husband has left on a dangerous mission. She has the choice of living either with his parents or hers, or of working and living with a girl friend.
a. What did she do and why?
b. How did she feel?
- (XII) 15. A girl has been feeling quite ill, but a doctor tells her she is not seriously sick and that it will do her no harm to go ahead and live a normal life.
a. What did she do and why?
b. How did she feel?
- qualified him for his chosen profession. He is offered a job but his parents suggest that he come home first for a rest.
a. What did he do and why?
b. How did he feel?
- (IV) 4. A young couple are in love, but she maintains that a difference in racial and religious background will make an unhappy marriage.
a. What did he do and why?
b. How did he feel?
- (V) 5. A young man is engaged to be married, but when he introduces the girl to his parents, one of them raises strong objections.
a. What did he do and why?
b. How did he feel?
- (XIII) 6. A young husband is able to support his wife. But she tells him she wants to keep on working because she prefers to be financially independent.
a. What did he do and why?
b. How did he feel?
- (XIV) 7. A young man cannot support his mother and his wife. The girl to whom he is engaged wants to be married.
a. What did he do and why?
b. How did he feel?
- (VI) 8. By posing as better trained and qualified than he is a young man has a chance at an excellent position. If he is frank about his background he will not get the job.
a. What did he do and why?
b. How did he feel?
- (VII) 9. A couple have been going together, but the young man notices that his girl friend is very cordial to a friend of his own who often accompanies them on dates.
a. What did he do and why?
b. How did he feel?
- (VIII) 10. A young man gets into trouble. If he tells his parents they can help him, but they will be much disappointed and will strongly disapprove of what he has done.
a. What did he do and why?
b. How did he feel?
- (XV) 11. In order for his wife to keep on with her job it is necessary for a young father to help with the housework and care for the children.
a. What did he do and why?
b. How did he feel?
- (IX) 12. A father offers to pay for his son's

FORM 2—Men

- (I) 1. A young man is much pursued by a girl who is attracted to him. She takes the initiative in suggesting dates.
a. What did he do and why?
b. How did he feel?
- (II) 2. A young man who has been very popular at home is teased by his present companions about his clothes and speech.
a. What did he do and why?
b. How did he feel?
- (III) 3. A young man fails an important examination which would have

education if the latter will take courses he prescribes. Otherwise the son must pay his own way.
a. What did he do and why?
b. How did he feel?

- (X) 13. A young man is engaged to a girl who wants him to change certain habits, manners and ways of acting.

a. What did he do and why?
b. How did he feel?

- (XI) 14. A young man notices that people tend to avoid him.

a. What did he do and why?
b. How did he feel?

- (XII) 15. A young man is told by a doctor that he is in poor health and needs a rest, but this will seriously interfere with his work.

a. What did he do and why?
b. How did he feel?

FORM 2—Women

- (XIII) 1. A girl has a job away from home. Her parents write her that a relative has left them some money and suggest that she give up the job and come home to live with them.

a. What did she do and why?
b. How did she feel?

- (I) 2. A girl finds that her friends get dates by playing up to the men and taking the initiative. They tell her she is slow.

a. What did she do and why?
b. How did she feel?

- (XIV) 3. A girl is very much attracted to a man she has known only three weeks. He wants her to marry him before he leaves on a dangerous mission. Her parents object.

a. What did she do and why?
b. How did she feel?

- (XV) 4. A girl is told by someone that an intimate friend of hers has said unfavorable things about her.

a. What did she do and why?
b. How did she feel?

- (III) 5. A girl fails an important examination which would have qualified her for her chosen profession. She is offered a job, but her parents suggest that she come home for a rest.

a. What did she do and why?
b. How did she feel?

- (II) 6. A girl whose parents are divorced is asked to testify against one in

a lawsuit involving her own inheritance.

a. What did she do and why?
b. How did she feel?

- (IV) 7. A young couple are in love, but he maintains that a difference in racial and religious background will make an unhappy marriage.

a. What did she do and why?
b. How did she feel?

- (V) 8. A girl is engaged to be married, but when she introduces the young man to her parents, her father raises strong objections.

a. What did she do and why?
b. How did she feel?

- (VI) 9. By posing as better trained and qualified than she is, a girl has a chance at an excellent position. If she is frank about her background she will not get the job.

a. What did she do and why?
b. How did she feel?

- (VII) 10. A couple has been going together, but the girl notices that her boy friend is very cordial to a friend of her own who often double dates with them.

a. What did she do and why?
b. How did she feel?

- (VIII) 11. A girl gets into trouble. If she tells her parents they can help her, but they will be much disappointed and will strongly disapprove of what she has done.

a. What did she do and why?
b. How did she feel?

- (X) 12. A girl is engaged to a man who wants her to change certain habits, manners and ways of acting.

a. What did she do and why?
b. How did she feel?

- (XI) 13. A girl notices that people tend to avoid her.

a. What did she do and why?
b. How did she feel?

- (IX) 14. A father offers to pay his daughter's tuition in school if she will take the courses he prescribes; otherwise she must earn her own way.

a. What did she do and why?
b. How did she feel?

- (XII) 15. A girl is told by a doctor that she is in poor health and needs a rest, but this will seriously interfere with her work.

a. What did she do and why?
b. How did she feel?

TABLE 12
Armatures distributed by area of conflict

Conflict area	Number of armatures*			
	Form 1		Form 2	
	Men	Women	Men	Women
Family	5	7	5	6
Opposite sex	7	6	6	5
Social and friendship	3	3	4	4
Religion and opinion	1	1	1	1
Vocation	3	3	2	2
Health	1	1	1	1

* The total of armatures in each area is more than the total number of items on each form, since some may be classified in more than one area.

TABLE 13
Category weights for Form 1 M

Armature	Categories*														
	I	2	3	4	5	6	7	8	9	10	El	Ev	Q	S	
I	3	3	1	4	3	3	3	2	3	3	3	3	3	3	
II	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
III	3	3	3	3	3	3	3	2	3	3	3	3	3	3	
IV	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
V	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
VI	3	3	3	3	3	3	3	2	3	3	3	3	3	3	
VII	3	3	3	3	3	3	3	3	2	3	3	3	3	3	
VIII	3	3	3	3	2	3	3	3	3	3	3	3	3	3	
IX	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
X	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
XI	3	3	3	2	3	3	3	3	3	3	3	3	3	3	
XII	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
XIII	3	3	3	3	3	3	3	3	3	1	3	3	3	3	
XIV	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
XV	3	3	3	3	3	1	3	3	3	3	3	3	3	3	

* See table 4, p. 12, for names of categories.

TABLE 14
Category weights for Form 1 W

Armature	Categories*														
	I	2	3	4	5	6	7	8	9	10	El	Ev	Q	S	
I	3	3	I	3	3	3	3	I	3	3	3	3	3	3	
II	3	3	2	3	3	3	3	3	3	3	3	3	3	3	
III	4	3	3	3	3	3	3	3	3	3	3	3	3	3	
IV	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
V	3	3	4	3	3	3	3	4	3	3	3	3	3	3	
VI	2	3	3	3	3	3	3	I	3	3	3	2	3	3	
VII	3	3	3	3	3	3	3	3	2	3	3	3	3	3	
VIII	3	3	3	I	I	4	2	3	3	3	3	3	3	3	
IX	3	3	I	3	2	3	3	3	3	3	3	4	3	3	
X	3	3	3	3	4	3	3	3	3	3	3	3	3	3	
XI	3	3	3	I	2	3	3	3	3	3	3	3	3	3	
XII	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
XIII	3	3	3	3	3	3	4	3	3	3	3	3	3	3	
XIV	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
XV	3	3	3	3	3	3	4	3	3	4	3	3	3	3	

* See table 4, p. 12, for names of categories.

TABLE 15
Category weights for Form 2 M

Armature	Categories*													
	1	2	3	4	5	6	7	8	9	10	El	Ev	Q	S
I	4	3	3	3	3	1	3	3	3	3	3	3	3	3
II	3	3	3	3	3	3	3	3	3	3	3	3	3	3
III	2	3	3	3	3	3	3	3	2	3	3	3	3	3
IV	3	3	3	3	3	3	3	3	3	3	3	3	3	3
V	3	3	3	3	3	3	3	3	4	3	3	3	3	3
VI-XV#	3	3	3	3	3	3	3	3	3	3	3	3	3	3

* See table 4, p. 12, for names of categories.

Frequencies too small for computation of weights. Of the nine NU men who were subjects for Form 2, all answered the first five questions. There was, however, little overlapping in their choice of additional armatures. The standard weight of 3 was used for the remainder of the armatures.

TABLE 16
Category weights for Form 2 F

Armature	Categories*													
	1	2	3	4	5	6	7	8	9	10	El	Ev	Q	S
I	3	3	3	3	3	3	3	3	3	3	3	3	3	3
II	3	3	4	3	3	4	3	3	3	4	3	3	3	3
III	5	3	3	2	1	3	3	3	3	3	3	3	4	3
IV	1	3	3	3	3	3	3	2	3	3	3	3	3	3
V	3	3	3	3	3	3	2	3	3	3	4	3	3	3
VI	3	3	3	3	4	3	3	5	3	3	3	3	3	3
VII	3	3	3	3	3	3	3	3	3	3	3	3	3	3
VIII	3	3	3	3	3	3	4	3	3	3	3	3	3	3
IX	3	3	3	3	3	3	2	3	4	3	3	3	3	3
X	3	3	3	3	3	2	3	3	3	3	3	3	3	3
XI	3	3	3	3	3	4	3	3	3	3	3	3	3	3
XII	3	3	3	3	3	3	3	3	3	3	3	3	3	3
XIII	3	3	3	3	4	3	3	3	3	1	1	3	1	3
XIV	1	3	3	4	3	3	3	3	2	3	3	3	3	3
XV	1	1	3	1	3	3	3	3	3	3	3	3	3	3

* See Table 4, p. 12, for names of categories.

APPENDIX B

INSTRUCTIONS TO JUDGES FOR RELIABILITY EXPERIMENTS

In the following section, the three separate sets of instructions which were given to the judges who took part in the reliability experiments are presented. There is, of course, considerable repetition and overlapping as well as the introduction of certain altered concepts in each new set. The instructions in their original form have been preserved intact for reference in connection with the results of the several experiments, and as a further description of the way in which the final scoring system evolved.

RELIABILITY EXPERIMENT I

Classification of feeling expressions

1. Expressions of *frustration* (frustration of external origin)
 - a. criticism, animosity, opposition
 - b. restraint, forcing, punishment
 - c. difficulty, lack, unpleasant situation
 - d. shock, loss, failure, surprise (unpleasant)
2. Expressions of *aggression* (aggressive action tendencies)
 - a. aggressive action, revenge, retaliation, reproach
 - b. self-forwarding, independence
 - c. persistence, determination
3. Expressions of *submission*
 - a. dependence, clinging
 - b. renunciation, sacrifice
 - c. acceptance, adjustment, resignation, detachment
4. Expressions of *avoidance*
 - a. evasion, retreat, regression, relaxation of effort
 - b. concealment, cover-up, hide feeling
 - c. indifference, forgetting
 - d. preserve harmony, act with caution
5. Expressions of *depressions* (attitude as opposed to event)
 - a. unhappiness, misery, grief, loneliness
 - b. disappointment, disillusionment, pessimism
6. Expressions of *pleasure* (includes both mood and event)
 - a. elation, expansion
 - b. aspiration, optimistic expectancy
 - c. confidence, satisfaction, pride
 - d. enjoyment, interest
7. Expressions of *negative attitudes*
 - a. criticism, rejection, dislike, hate
 - b. resentment, outrage, disgust
 - c. resistance, opposition
8. Expressions of *positive attitudes*
 - a. attraction (sex)
 - b. affection (filial, friendship)
 - c. tolerance, forgiveness
 - d. sympathy, pity
 - e. service, support, "nurturance"
 - f. cooperation, sharing
9. Expressions of *anxiety and conflict*
 - a. wish, desire, longing
 - b. effort to overcome defect, self criticism
 - c. embarrassment, inadequacy, inferiority, shyness
 - d. fear, worry, dread
 - e. remorse, regret, shame, guilt
 - f. uncertainty, doubt, conflict, confusion
 - g. duty, obligation, atonement
10. Expressions of *rationalized feeling*
 - a. moralizing
 - b. self justification
 - c. evaluation

Instructions to judges

As one approach to scoring the projective "Test of Insight into Human Motives" responses written by subjects have been broken up into phrases which appear to express some kind or degree of feeling. Results so far have shown that subjects differ not only in the amount of feeling expressed in proportion to the number of lines written, but also in types of affective expression most frequently used.

On the basis of kinds of feeling expression appearing in fourteen papers (written by six men and eight women) the attached set of categories has been set up empirically. In order to check on the reliability of the classification, it seems advisable to have it tried out by four judges in addition to the experimenter, for the purpose of computing percentage of agreement.

On each of the slips in the package is a word, phrase, or part of a sentence. An attempt has been made to include enough of the original sentence to make the intent of the phrase discernible. Whole sentences have not been presented, since several different feelings are often present in a single sentence, and for our purpose it is desirable to isolate them. Although in the original responses the

phrases were written concerning fictitious third parties, "he" or "she" is to be taken as referring to the writer. Unless otherwise indicated, such phrases as "felt miserable," "was happy," also refer to the subject. Where the reference is not clear, "S" has been added in parenthesis to indicate the subject. References to other than the subject are obvious by the context (inclusion of words such as "parents," "friend," etc.). Words and phrases in parenthesis on some of the slips are not part of the original response, but were added to make the meanings more evident.

As one of the judges you are asked to proceed as follows:

1. Sort the cards into the ten main categories, using the subheads as a guide as to what the categories are intended to include.
2. Sort the main categories into their subdivisions.
3. Enter the number and letter designation of the category on each card. It will probably be necessary to do some experimental sorting and reclassifying before entering your final judgment. If there are any which you cannot classify, please designate by X.
4. If the category scheme as presented does not seem satisfactory to you, please devise another, using not more than ten divisions. Mark these also on the cards, using a different set of category numbers, such as 21-30, to distinguish it from this which uses 1-10.

Note: Steps 1, 2, and 3 are essential. Step 4 is optional.

RELIABILITY EXPERIMENT II

Classification of feeling expressions

Explanation to judges. As one approach to scoring the projective "Test of Insight into Human Motives," responses written by the subjects have been broken up into phrases which appear to express some kind of degree of feeling. Results so far have shown that subjects differ not only in the amount of feeling expressed in proportion to number of lines written, but also in types of affective expression most frequently used.

Four hundred and eight such phrases have been classified by four psychologists in addition to the experimenter. Percentage of agreement between all ten pairs of judges has yielded contingency coefficients which

show a substantial amount of agreement among all the judges, but a rather wide discrepancy between different pairings. Study of the individual use of the classificatory scheme indicates that the categories were not sufficiently well-defined, and that the judges were using different interpretations of the categories themselves. It is obvious that very high reliability cannot be established for any measuring scheme unless the same scheme is being used by all. Comparable results cannot be obtained when one surveyor uses, for example, a 14 inch foot-rule while another uses a 12 inch. It seems desirable, therefore, to train judges in the interpretation of the categories in order to establish uniform concepts before independent sorting of the items is undertaken.

It should also be borne in mind that the question of *reliability* of the category scheme is the only problem with which the present aspect of the experiment is concerned. Its *validity*—that is, whether or not it is a "good" system which is capable of evaluating differences between individuals and groups—must be established by other methods. Judges are asked to maintain a critical attitude toward the definition and explanation of the categories, since it will be desirable to express their meaning as clearly as possible. Suggestions on this point are sought. As far as the scheme as a whole is concerned, the judges are asked to take it on faith, until its validity can be established, and to concentrate attention on mastering it as it stands.

On each of the slips in the package is a word, phrase, or part of a sentence. An attempt has been made to include enough of the original sentence to make the intent of the phrase discernible. Whole sentences have not been presented, since several different feelings are often present in a single statement, and for our purpose it is desirable to isolate them. Although in the original responses the phrases were written concerning fictitious third parties, "he," or "she" is to be taken here as referring to the writer. Unless otherwise indicated, such phrases as "felt miserable," "was happy" also refer to the subject. Where the reference is not clear, "S" has been added in parenthesis to indicate the subject. References to other than the subject are usually obvious by the context (inclusion of words such as "parents,"

"friend," etc.). Words and phrases in parenthesis on some of the slips are not part of the original response, but were added to make the context more evident.

Instructions. You are asked to follow carefully the steps outlined below:

1. Study thoroughly the description of the categories on the ensuing pages. For your convenience a single page with all the main categories listed has been included. After you have mastered the system it will probably be easier simply to refer to this, but whenever a doubt arises you should refer to the full description.

2. Sort the cards into ten piles, using the list of subheads and the category explanations as a guide, always determining the proper classification of an item according to its congruence with the main category.

3. After your original sorting, go over each pile to determine whether any items seem out of place with others which have been included. In order to avoid shifting concepts as you proceed, experimental sorting is necessary before entering your final judgment.

4. Only after the first three steps have been taken should you write the number of the category on the face of each slip. Any which you cannot classify designate by X. (See also general discussion of the category scheme below.)

5. Return the cards to their original numerical order after recording your judgment.

CATEGORIES

General plan. The ten categories on the list fall into groups as follows: No. 1, "frustration," refers to the appraisal of a *situation* or *event* external to the subject. Nos. 2, 3, and 4, "aggression," "submission," "avoidance," all refer to that which a subject *does* about a situation. Nos. 5 and 6, "depression," and "satisfaction," or "pleasure," refer to mood, and Nos. 7 and 8, "positive" and "negative" attitudes, refer to attitudes toward other people or toward situations. The latter two might as well be termed rejection-acceptance or withdrawal-approach, but the terms "positive" and "negative" attitudes have been chosen to represent both. Nos. 9 and 10, "anxiety" and "rationalization," like No. 1, are not part of a pair or group of other categories.

From the above general description it can be seen that certain items could conceivably

be placed in several categories. No item properly classified could, however, fall in more than one category within a group. (One cannot aggress, submit, and avoid in a single act as defined below.) Likewise one cannot be both elated and depressed, or take both a positive and negative attitude, except when there is ambivalence and conflict which is provided for in category No. 9. It would be possible, however, for a single phrase to describe a frustration, to provide a hint as to what the subject does about it, to express a mood, and a positive or negative attitude. It might also indicate anxiety or rationalization. In most instances it is possible to decide which aspect of the phrase is paramount. When the judge is unable to decide this point, all choices of placement should be listed, and the one which seems most important should be encircled. This problem makes more trouble in items taken out of context than it does in actual test scoring. In the latter situation it is possible to divide up the sentence so that all important aspects of the feeling are counted. In classifying the separate phrases, it is best to take as your point of reference *what the words actually say as presented*, excluding as far as possible the implications which the judge will naturally imagine as part of the original setting.

1. Expressions of frustration:

All references to frustrating situations, whether personal (such as parental domination) or relatively impersonal (such as difficulty, lack of money, etc.). This category does not include the subject's feeling about the frustration, or his action concerning it, but only his mention of it. The category includes the following "frustrations" encountered by a subject, plus any number of others not specifically cited below. (The same will be true of examples given under the other categories.)

Examples: a. criticism, animosity, opposition; b. restraint, forcing, punishment; c. difficulty, lack, unpleasant situations; d. shock, loss, failure, surprise (unpleasant).

2. Expressions of aggressive action tendencies.

This refers solely to mentions of action taken or to be taken in a situation. It includes any description of the subject in the act of doing something to change a situation, or to bring about some end. Thus it covers aggression in the sense of attack or revenge.

and also aggression in the sense of active carrying out of some ambition, domination or persuasion of others, etc.

Examples: a. revenge, retaliation, reproach; b. self-forwarding, effort toward independence; c. persistence, determination.

3. Expressions of submission and acceptance.

Items belong in this category when the subject, instead of aggressing or actively avoiding a situation, submits to it or accepts it. The category may be thought of as opposed to No. 2. It includes both the extremes of dependence on others, and detached attitudes, e. g. when a person chooses to adjust to a situation rather than to adjust the situation to himself.

Examples: a. dependence; b. renunciation, sacrifice, letting go; c. acceptance, adjustment, resignation, objective detachment.

4. Expressions of avoidance:

Like No. 2, this category indicates that the subject does something about a situation, but here the tendency is to do something to avoid or escape, rather than to alter it. It includes expressions in which the subject is described as acting to preserve a *status quo* (thus avoiding some threat). It also includes various defenses and protective devices such as covering up, forgetting, cultivated indifference.

Examples: a. evasion, retreat, regression, relaxation of effort; b. concealment, cover-up, hide feeling; c. indifference, forgetting; d. preservation of harmony, action with caution.

5. Expressions of depression.

This includes various forms of unhappiness, disappointment, loneliness, disillusionment. Note the distinction between this and No. 1 (frustration). No. 1 mentions an external situation or event, No. 5 refers to the subjective reactions to such an event. For example, "He let her down" would be a frustration if the pronoun refers to "her" as the subject, but "she felt let down" would be classed under depression.

Examples: a. unhappiness, misery, grief, loneliness; b. disappointment, disillusionment, pessimism, bitterness.

6. Expressions of satisfaction and pleasure.

Since no category has been set up to cover pleasant events (antithesis of No. 1, frustrations) mention of them may be included

here. But note that this category is chiefly concerned with mood, as is No. 5, and that if action is described, the item should be referred to No. 2 (aggression). If the phrase expresses an effort to win happiness, it is action. If it merely refers to happiness, it is mood and belongs in this category.

Examples: a. elation; b. expansive feelings; c. confidence, satisfaction, pride; d. aspiration, optimistic expectancy, anticipation; e. enjoyment, interest.

7. Expressions of negative attitudes:

The line drawn between No. 7 (negative attitudes) and No. 5 (depression), and between No. 8 (positive attitudes) and No. 6 (satisfaction) is necessarily a fine one, but if the following distinctions are kept in mind it should be possible to distinguish them. A "mood" may best be considered as a subjective feeling tone, and an "attitude" as a "set" for action toward or away from something or someone. Obviously attitudes have feeling tones, and moods, condition sets, but for the present purpose the distinction will serve. For example, the phrase "he felt disappointed" describes a mood only, provided no other clues are included. Therefore it is classified under "depression." On the other hand, if a person is described as feeling disgusted or resentful, he may be depressed, but an attitude of active rejection is also implied. The following rule should be remembered: if action is mentioned, the item goes under No. 2; if some attitude which includes an action tendency is implied it goes under No. 7 or No. 8 according to whether it is positive or negative; if only mood is expressed, it belongs under No. 5 or No. 6, according to whether it is depression or satisfaction.

Examples: a. critical attitude, rejection, dislike, hate; b. resentment, outrage, disgust; c. resistance.

8. Expressions of positive attitudes.

Note that the same distinction should be made between No. 8 and No. 6, as has already been discussed in connection with No. 7 and No. 5. All of the more "socialized" attitudes may be classified here, such as liking or love, and other positive, approaching attitudes toward others.

Examples: a. sex attraction; b. affection (filial or friendship); c. tolerance, forgiveness, understanding; d. sympathy, pity; c. service, cooperation, sharing.

9. Expressions of anxiety and conflict.

This category includes a large number of items which are all characterized by dissatisfaction of some kind. Its chief distinction from No. 5 is that although all of the feelings included involve some form of unhappiness, they contain the added elements of *dissatisfaction with self*, and various forms of *conflict*. There is some question as to whether "wish, desire, longing" should be included here, rather than under "depression." It has been placed in this category because a wish usually expresses conflicting feeling, including "hope" (which is not depression) and "dissatisfaction" (which may be). "Fear" has also been included, because fear usually involves a sense of the inadequacy of the self to meet a threat, as well as forthright dread of something external.

Examples: a. wish, desire, longing; b. effort to overcome fault or personal defect, self-criticism; c. embarrassment, inadequacy, inferiority feeling, shyness; d. fear, dread, worry; e. regret, remorse, shame, guilt; f. uncertainty, doubt, conflict, confusion; g. sense of duty or obligation atonement.

10. Expressions of rationalized feeling.

This category includes certain items which cannot be properly placed in any of the others, yet care should be exercised not to use it as a dump heap. Even if an element of rationalization appears, if the feeling being rationalized is clear enough for classification in another category, this should be given preference. For instance, "he felt he owed his parents something" may be a rationalization, but it also corresponds to No. 9 g. and should be thus classified. No. 10 is to be used for expressions the meaning of which is not clear *except* as moralizing, philosophizing, self-justification, evaluation, and the like. Summary of category list for ready reference:

1. Expressions of frustration
2. Expressions of aggression
3. Expressions of submission and acceptance
4. Expressions of avoidance and escape
5. Expressions of depression
6. Expressions of satisfaction and pleasure
7. Expressions of negative attitudes
8. Expressions of positive attitudes
9. Expressions of anxiety and conflict
10. Expressions of rationalized feeling.

See above for fuller explanation.

RELIABILITY EXPERIMENT IV

Scoring instructions

(Enter scores on the back of each answer)

Column 1. Feelings

1. List categories by number and mark the number of feelings in each category.⁵⁴
2. Add up total number of feeling expressions.
3. Divide total number of feeling expressions by number of typed lines to get feeling ratings. (Note: count as a line any line that contains more than two words.)

Column 2. Miscellaneous indicators

1. Elaboration. Two types will be found: (1) Subject adds something to the situation which is not stated or implied in the armature. (2) Subject mentions some event which happens to the character *following* the situation, usually as a result of what he had said or felt.
2. Evaluations. Feeling category No. 10 takes care of some of these, when the "evaluation" or "generalization" is a rationalization of something the character does or feels. Other such statements are inserted without direct reference to the action described. Both are to be counted and listed in the evaluation total.
3. Use of first person pronouns, "I," "me," etc.
4. Qualifications. Such expressions as "it would depend," "that is, if . . ." Any statement in which a subject mentions an action or a feeling, and then backs away from it, is counted as qualification. This does not include those answers in which two different solutions are stated, each based on a different set of premises.
5. Subjectivisms. These include all statements which cannot be interpreted without knowledge of their meaning to the subject. The category covers the senseless material inserted by some of the mental patients, but

⁵⁴ The following change has been made in the categories: No. 2 (aggressive action) has been eliminated. Feelings in this category can be placed in No. 7 (negative attitudes), No. 8 positive attitudes) or No. 9 (anxiety). The "action" part of the category is taken care of by the classification of solutions (see below). In its place is a new No. 2 "irrelevant" feeling, which you may find necessary for scoring the answers of some of the mental patients, and possibly a few of the normal subjects.

is sometimes found in the normal papers.

6. Answers without content (AWC). Some of the mental patients write elaborate statements explaining why they cannot answer the questions, or say nothing which has bearing on the armature, and nothing which can be classified in one of the feeling categories.

7. Suspicion. Expressions of suspicion, when they are attributed to the characters, may be scored as No. 9, one type of anxiety. When very marked, however, or when referred to the question or to the test situation itself, it should be noted separately in column 2.

Column 3. Types of conflict solution

1. Attack. The character does something active. He approaches, takes advantage of, says or does something to alter the situation, etc.

2. Resistance. The character tries to maintain the *status quo*, resists the infringement of something new which might upset his chosen way.

3. Acceptance. The character accepts the situation as it is. He may feel disturbed, but he does nothing different, does not shift his course.

4. Compromise. The character finds some

way to get around the situation, without directly meeting it, and without entirely changing his own previous course.

5. Escape. This includes retreat, avoidance, and the like. Any action which is designed to or actually does accomplish these ends is to be included here.

6. Conflict, confusion. The character tries various ways to solve the problem, without success. Such answers contain mostly feeling expression and little action.

7. Overcompensation. The character goes to extremes in whatever action the situation demands.

8. Reaction formation. The character goes to the opposite extreme, substituting some other act for one which the situation directly requires.

9. Unreal solutions. The situation is solved by an elaboration; that is, the subject introduces some element which is not present in the armature, in order to make the situation "turn out right."

10. Exploration and investigation. The character's approach to the situation is to find out more about it before going ahead, or to use it to gain other types of information or experience.

APPENDIX C

EXAMPLES OF SCORING AND INTERPRETATION

The manner in which the scoring system was applied is illustrated by the following exhibits:

1. A scored answer to question III, F2W, showing the notations used, and the method of summarizing the various factors.

2. The full paper written by subject EM4, chosen not because it is of more striking interest than other papers, but because it is of suitable length and demonstrates a variety of scoring factors.

3. A sample of the chart used for summarizing individual scores and converting them into profiles, drawn up for subject EM4.

4. An interpretive sketch for this subject, based on both quantitative and qualitative factors and presented in the form developed for individual analysis.

SAMPLES OF ANSWER SCORING

2W-7-III⁵⁵

- A. The girl takes the job, because/she knows she'd never feel quite right with herself again/(if she gave up at the crucial point.) /She is terribly disappointed/at first about her failure, (bu) knows her years of training cannot be judged in one exam/So she takes the job/knowing she'll then continue studying till she makes good./and/will not have time/to/sit and condemn herself/and/feel mistreated/if she goes home, and sits around brooding.)
- B. /She feels somewhat disappointed in herself/(but knows)by getting to work at something she can better alleviate her disappointment,xgain a new perspective as she keeps herself busy./She will come to have a sense of self-satisfaction/after, in the fact that she faced her problem instead of trying to escape from it, and withxthis healthy attitudexwill probably go on to eventually succeed in her chosen field.)

⁵⁵ The notation indicates women's Form 2 (2F), NU subject 7, armature III.

TABULATION OF SCORES FOR ANSWER 2W-7-III

Feeling	Elaboration	2	Solution
4. 1	Evaluation	2	Attack
5. 2	Qualification	4	
6. 2			
9. 2			
10. 1			

Key to scoring notations:

Underlines: conflict solution (action taken)

Brackets: < > : feeling expressions

Brackets: \ / : elaboration

Brackets: x x : evaluation

Brackets: C) : qualification

Numbers 1-10; feeling categories (See table 4, p. 12)

COMPLETE PAPER

WRITTEN BY SUBJECT EM4 (FORM 2)

I

I would accept the dates if I liked the girl well enough to go out with her. The reason is a gentleman would not embarrass a lady by not accepting her.

I would feel quite flattered to be asked to go out with her.

II

I would try to find out what it is that is wrong and try to rectify my mistakes because I would like to please my parents and my friends both.

I wouldn't feel very bad about it as I think they would be doing me a favor so I can change my way of dressing and speech.

III

I would accept the position because it would relieve the aggravation of failure in my chosen profession.

I would feel rather bad about it but would not let it interfere with my present position.

IV

I would try and convince the girl not to feel that racial and religious background would make an unhappy marriage as I think that if two people love each other it would not make a great deal of difference.

SUBJECT: EM4

ANSWER ANALYSIS CHART

5
8

$$\frac{F_{\text{TOTAL}}(\#1+\#3-10)}{NA} = 6.86 \text{ } 1$$

$$\frac{C_{TOTAL}(E_I + E_V + Q)}{NA} = 2.60 \text{ z}$$

$$\frac{\text{II TOTAL} (\#2 + P + S)}{\text{NA}}$$

F RANGE = 1.20 2

C RANGE = .20 2

I Range = 0 1

ZERO CATEGORIES: 1, 2

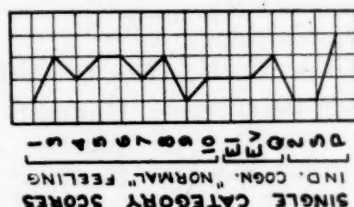
ZERO ARMATURES: 0 1

F CONVERTED: 31.3

C CONVERTED 41.6

$$F = .76$$

PROFILES



QUARTILES 1 2 3 4



QUARTILES 1 2 3 4

[illegible]

CONFLICT SOLUTIONS

AT RE AC CP EX

KEY TO CHART

R: RAW SCORE

W = WEIGHT (3 UNLESS INDICATED)

S = WEIGHTED SCORE

FIG. 7. Sample of answer analysis chart used for summary and interpretation of individual papers.

V

I would try and convince the objector that this girl is the right one for me because I think I would know what is best for me even if the objector still does not see my point of view I would feel as though I am old enough to know my right mind and would pay no attention to it.

XIII

I would try and convince my wife that I would prefer her not to work because a woman has enough to do at home. I would not like it as it would make me feel as though I could not take the care of her.

XIV

I believe I would support my mother before I would get married as I think I owe more to my mother than I do to my happiness. Until I would be able to support both. I would feel rather bad about it for a time.

VI

I think I would take a chance at posing as being able to handle the job because most times you can learn enough when you start to handle it efficiently very soon after. I feel as though I would not expose my background unless it was absolutely necessary. If not I would try and get a job I could qualify for.

VII

I would try and show more consideration and affection to my girl friend to try and distract her attention from my friend or else discontinue going out with him. I would feel rather down hearted that she prefers him to me.

VIII

I would tell my parents because the parents would be understanding and I think they feel they want to help you as much as possible, and most always want you to confide in them.

I would feel very much ashamed of myself and would try and make it up to my folks for helping me out.

XV

If it were necessary for my wife to work I would help take care of the children and

the housework because it would greatly be appreciated to know my wife wants to help provide for my home.

IX

I would prefer paying my own way if I had to take up something I didn't care about because I think I would know what I would like to do best.

I would feel as though my father thought I didn't have enough good judgment to pick out my own course.

X

I would change those certain ways of acting and habits if I thought they were best for me, otherwise not.

I would feel that my wife is trying to be very domineering and it would not appeal to me.

XI

I would make it my business to find out why people avoid me and try and rectify it. I would feel quite embarrassed.⁵⁶

XII

I would take the rest because I believe my health is more important than my work no matter how important my work may be, you can do much better when you have had your rest than if you try and continue while you are ill. I would take the rest good naturedly as worry would not help it any and you would recuperate⁵⁶ much quicker.

INTERPRETIVE SKETCH

Subject: EM4

I. Profile

A. Feeling expression

1. Level: low
2. Range: narrow average
3. Emphasis: depression, dependence, positive feeling.

B. Cognitive expressions

1. Level: low average
2. Range: narrow average
3. Emphasis: some elaboration and evaluation; greater amount of qualification.

C. Maladjustment indicators:

1. Level: high

⁵⁶ The misspelling is the subject's.

2. Range: narrow

3. Emphasis: score based entirely on excessive number of first person pronouns.

D. Conflict solutions: almost equal emphasis on attacking and accepting solutions. No "unreal" and no "no solution" answers

E. F/C quotient: .76 (average for E, low for NU)

II. General impression based on profile

Favorable signs in this record are the ability to express positive feeling, the fact that the number of unscored categories is average for normals, and also the fact that the feeling range has almost an average spread. The absence of irrelevant feeling, subjectivism, and "unreal" and "no solution" answers is also encouraging. Except for the excessive number of first person pronouns, this might be the paper of a poorly educated normal individual.

III. Additional impressions after reading paper

This individual seems extremely egocentric in his approach, discussing everything in terms of what he himself would do. There is considerable evidence of a strong parental

tie (see references to wish to please parents, choice of mother in preference to wife, "I owe my mother more than I owe my happiness," and "the parents would be understanding,") which is sometimes indicative of the dependent personality. There are a number of statements of a depressive type, such as "I would feel bad about it," and the like. The general overall impression is of an inadequate individual, probably more neurotic than psychotic, since the low affect sign is somewhat counteracted by favorable indications.

NOTE: The hospital record shows that the patient is a chronic alcoholic with suicidal tendencies. He was self-committed to the institution several times, and at the last notation was being considered as an attendant helper. Within the institution he makes a very satisfactory adjustment, but appears unable to meet the situation outside. His first breakdown followed desertion by his wife. It is interesting to note that the test record independently supported three major facts about the case: (1) absence of psychosis, (2) dependent personality with potentialities for rapport (positive feeling expression), and (3) tendencies toward depression.

APPENDIX D

REPLIES OF SUBJECTS TO VERBAL QUESTIONNAIRE⁵⁷

Question 1. "How did you go about solving the situations on the test?"

Question 2. "Did you use your own experiences to help you figure out what the characters would do, or did you just make them up?" (Question 2 was used only if information was not brought out in reply to question 1.)

Subject	Answer
W 1	I said what I would do. . . . I think you answer most questions according to how you've been brought up.
W 3	I thought you wanted what I myself would do. . . . Some were pretty far out of my experience. . . . I couldn't feature myself in their position.
W 4	I read through, picked the ones that applied to me and so tried to think what they would do. I used my own or other people I've known who have had the same experience.
W 5	Well, I read them over first and then kind of thought. I didn't put myself into any of them. I placed myself outside the situation. My first impression was what I went on because otherwise I would think of alternatives and just write on and on. (Q.) Well, maybe in the one I just mentioned. Most of the others were beyond anything I've experienced.
W 6	I thought—I didn't take the one case—well—my attitude was what most people would do. I couldn't see myself in most of them. . . . Well, the first thing I thought of was what I would do. But it might be the opposite.
W 7	Well, I did it different from most people. Some people would put how someone else would act. I used my own reactions. It was purely personal.
W 8	First I tried to think of a case, such

⁵⁷ The questions were not asked in rigid order but were used informally with phrasing somewhat varied. The replies given are not full reports of everything said, as differing amounts of personal information were volunteered and encouraged, interspersed between answers. The above replies, although excerpted for placement with the question answered, are from nearly verbatim notes.

as when such a thing had happened to me (I'm kind of selfish) and if I couldn't then I would try to think of a hypothetical case, yes, I did use my own experiences whenever I could. I don't see how you can do it any other way.

W 9	Well, I tried to think of what I would do if I were in the situation. Most of them were outside my experience.
W 10	I found it a little difficult to think myself back to that stage, but I tried to do it and to use my own experience and have the people respond as I would.
W 11	I think I did it differently in each batch. The second time I tried to think of a situation and go into it more. The first time I tended to quick judgments, my first reaction.
W 13	First I tried to write down what was nearest to my own situation. I wrote down what I thought. They were things I had thought about before.
W 14	I think I sort of put myself in the position of the girl and had her do what I would do. I mostly imagined.
W 15	Well, I was a little confused—didn't know how much of a supposition I was supposed to make, so I just gave my personal opinions. I used my own experiences and those of friends of mine.
W 16	Well, I tried to use my own experience. Of course there were some I just couldn't imagine happening to me. Like the one about the parents—I never had that happen.
W 17	It gave so little information I thought I'd put what most people would do—myself and my friends. . . . I used my own experiences and in a couple I put the opposite of what I would do.
M 1	Well, I answered just from the experiences of other people. (Q.) Probably, to some extent, in a few cases, or tried to use my own reactions, what that would be, I mean.
M 2	I imagined some sort of indifferent person—as if you'd given me the plot of a story and asked me to write the end. I tried to think what he would do. . . . I used my own experience in one of them—I don't mean my

clothes have been so funny I've been teased, but if they did criticize me, well, I used my own beliefs in that one—I think if a boy's friends criticized him he would feel he oughta see if he can't do something about it—take it. He ought to try to correct it. There are lots I don't know what I'd do. If I was in love with a girl who was Catholic or something, I don't know what I'd do about it.

M 3 Well, first I thought what I would do—what most people would do. Y—es, most of them I could imagine.

M 6 I figured out what I would do. Used things I've known and things that happened in the past. Most of them seemed to apply to me and people I know.

M 7 I think I answered a great deal as I would have reacted.

Summary of Questions 1 and 2.

Projection of own experiences and attitudes	11
Used own and also others' experiences	7
Denied use of own experiences	2

Question 3. "Which of the questions do you recall best now? Did any of them especially interest you? Were some easier or harder to work out than the others?"

Subject	Answer
W 1	I remember the one about the girls going out Saturday night and should they date and do what their friends do.
W 3	I liked those about the parents objecting—perhaps because I'd thought about it more—so many people my age are having that trouble.
W 4	I remember best the young lady who had just lost her chance—should she go home. I guess that was easiest because I've had the same problem—Well, will have. When I get through—going home to the family, or whether to stay and get a job.
W 5	I remember best the one about the girl's girl friends who thought she wasn't getting dates because she didn't go after men. I think that's a commoner experience. I've seen it oftener in others, I guess.
W 6	I can't remember—well, in one the male was being cordial to the other girl, and the one about taking subjects her father prescribes.

W 7 I remember the one where the girl wanted to get married and the boy was going on a dangerous mission and the parents objected. And the girl who found someone talking about her and didn't know what to do. Then there was one about the girl who did poorly in college. I suppose because I'm interested I wrote on the marriage ones. Lots of girls I know are getting married and its sort of a rabid subject. I think I fixed it so she waited. I had her talk it out with him.

W 8 I remember the one where the girl's boy friend wanted her to change something and the one where the mother wanted her to come home. The first I remember because it had personal reference; the second I couldn't imagine happening to me.

W 9 I remember best that one about inter-racial marriage. You see I lived in South America and we were always hearing about it.⁵⁸

W 10 I remember best the one about the girl who was not sure of her professional experience and whether she would choose a career—there were several. I don't just remember how they went but the career ones are personally more interesting. That matter of marriage and all was less personal to me.

W 11 I remember the girl discussing political views with her parents. I guess because that's happened in our family—my sister. Then the one about the army wife whose husband went to war.

W 13 I remember I didn't answer the ones about religious questions. We just don't usually discuss those things. I remember the one about the girl who was criticized by her friends for not dating. I've thought about that a great deal.

W 14 The easiest was the girl who worked away from home. Of course she would go back.

W 15 I remember the one where the father wanted the daughter to take certain courses. I guess that was because I couldn't imagine my Dad doing that.

⁵⁸ A follow-up report on this subject mentions that the girl has been going out with a young man, but that there is a religious conflict with parental disapproval. It is interesting to note the personal significance of this recall, and the rationalization given to account for it.

W 16 The first stayed in my mind—maybe because it was first—the girl living away from home. My parents wouldn't do anything like that. They think you *should* live away from home. (Oh! yes! I do too. I've learned more!) On some of them I could have written pages. Then others—especially about falling in love with a man of a different race—I can't imagine that.

W 17 I liked the one best about the one who was engaged and the man wanted her to change.

M 1 I recall the one about the fellow and his job away from home and the one about the fellow engaged to a girl of different religious beliefs. And the one about the fellow who came home with a lot of views he'd acquired away. I'd thought more about those things.

M 2 (See answer of this subject to Questions 1 and 2)

M 3 Denies any special recall. Cannot repeat any of the situations.

M 6 No memories.

M 7 There was one question I told all my friends about—the girl chasing after the young man. Then, that one about the man people avoided. I couldn't help thinking about the ads about soap but that was too obvious. I didn't think that was what you meant, so I stayed clear of that. I only answered five. There was one about the girl chasing the young man. And the one about the young man having to support his mother. That interested me. When I first looked it over I could have taken almost any one. But—I was looking for novel situations.

Summary of Question 3

Spontaneous admission of personal significance	12
Memory of specific items, but personal reference not volunteered	6
No memories	2

Question 4. "How did you approach the problems? Did you say what you thought the people *would* do, or what you thought *they* should do—what it would be right for them to do?"

Subject	Answer
W 1	You'd have to know the situation

before you could say what they *should* do. Here you don't know enough about it. You just have to imagine.

W 3 Would, I guess, but I was quite practical. I thought if I were in that position I would do what was right. So—I guess I had them do as they should. Both *should* and *would*, I guess.

W 4 I sort of felt what I would do. No, I didn't try to have them do the right thing.

W 5 I just put what I thought they *would* do. Of course you could go on for hours on what they should.

W 6 What they would do—based on what I've seen.

W 7 I put what they *would* do.

W 8 Would, rather than *should*. A lot of times when something happens, when we encounter a situation, we don't always know what we *should*. I thought that was the logical way to go about them.

W 9 I think I put what they would. In some of them I had a hard time deciding.

W 10 In most cases what they *would* do, I think

W 11 Well, now let me think. I think sometimes *should* and sometimes *would*. I think the first time I said what they would do, and the second time I went into it more, saying what they should.

W 13 I think I put what they *should* do.

W 14 I put what I thought they *should* do.

W 15 What they *should* do.

W 16 I'm sure I said that I thought they *should*.

W 17 What they *would* do.

M 1 I don't know—what they *would* do I think—not what they should.

M 2 One reason I didn't enjoy the test so much was because I didn't know whether to put *should* or *would*. I think they should have said. I used *would*, but that's why it's sort of confusing.

M 3 What they *should* do.

M 6 What they *would* do.

M 7 I think I put down what they would do—my immediate reaction. I was thinking of giving two alternatives, but I put down only the first.

Summary of Question 4

Told what the characters <i>would</i> do	12
Told what the characters <i>should</i> do	6

Used both *should* and *would* 2

Question 5. "Do you like to do this sort of thing? Was it sort of a bore, or did you find it easy, or hard?"

Subject

Answer

- W 1 Well, a lot of them were the kind you read in *Collier's* or the *Saturday Evening Post*, advice to the lovelorn or something—seemed a funny kind of thing to be coming out of a psychology department. . . . Some applied very well to things girls get together and talk about, like those on religion. The first hour was interesting, but when I came to the second, oh, well, that was interesting too. The questions were different, indefinite and vague. It's really a test of imagination, I think, unless you use yourself.
- W 4 No, I liked doing it. It gave me a chance to get my own thoughts collected on things I hadn't thought much about—sort of as vague possibilities, but not immediate. I enjoyed doing it.
- W 5 I think it would be more interesting if you weren't putting it on paper—analysing people, I mean. I think it would be more interesting if I knew the person more—if you had more to judge. After all, this didn't give you much to go on—I think I'd like it, though.
- W 6 I thought it was so ambiguous—everyone one would do something else. How can you tell?
- W 7 Oh, I love that sort of thing. I love to take psychological tests. I like to put my ideas down.
- W 8 Oh, I liked them. I think it was fun to take. The Insight one was very vague—it didn't give you much to go on. Yes, I think most of the situations are of everyday occurrence.
- W 9 I liked it when it was the kind I could do, something I knew. The others I found hard.
- W 10 I enjoyed it. As I say, it's all new to me. I don't know as I'd thought in

that trend. Of course, I don't think anything related to human behavior is uninteresting—if concerned with why people do as they do.

- W 11 At first I was quite excited about it, it was lots of fun. Then it began to get a little boring. That's why I picked fewer questions the next time, because some of them didn't interest me a bit.
- W 13 It was interesting enough to take, but I tried to figure out how you could use it. It seemed almost useless because it was so varied.
- W 14 I thought it was interesting to see what I thought they would do.
- W 15 I liked it. I thought it was very interesting.
- W 16 I thought it was interesting. It made you think about things you hadn't before. We all discussed it on the way back.
- W 17 I liked it but there was so little information, I think I felt kind of lost.
- M 1 It kind of got me wondering about it and about the purpose. There were so many different things they could do it hardly seemed worth doing.
- M 2 I think the personality test had more on the ball. I got more out of it. The first (Insight test) was interesting but it doesn't do too much good. . . . I like a question where I have some beliefs or experience to go on—like that one about giving up college to go home. I don't know *what* I'd do.
- M 3 Oh, I think it was interesting. My answers were not very long, but I liked it.
- M 6 I thought it was fairly interesting. I wrote right on through. I read them all, and then I took the first and went through.
- M 7 I liked it very much. I enjoyed doing it a lot.

Summary of Question 5

Enthusiastic about the test	1
Liked it quite well	8
Liked it fairly well	6
Mildly disliked	4
Strongly objected	1

APPENDIX E

SUPPLEMENTARY TABLES

TABLE 17
Reliability experiment I: Percent of agreement and contingency coefficients

		Judge A	Judge B	Judge C	Judge D
Judge B	%	50.2			
	C	.812			
	S.E.	.017			
Judge C	%	50.0	44.3		
	C	.800	.752		
	S.E.	.015	.024		
Judge D	%	46.0	41.9	41.4	
	C	.768	.730	.725	
	S.E.	.022	.026	.027	
Judge E	%	25.8	28.9	25.4	26.7
	C	.465	.536	.460	.488
	S.E.	.050	.045	.051	.048

t:t' = 10:408, n = 5, N = 2040

Percent of agreement for average judge = 38.5

Contingency coefficient for average judge = .68 ± .032

TABLE 18
Reliability experiment II: Percent of agreement and contingency coefficients

		Judge A	Judge F	Judge G	Judge H
Judge F	%	54.5			
	C	.830			
	S.E.	.021			
Judge G	%	65.0	52.5		
	C	.880	.818		
	S.E.	.016	.023		
Judge H	%	54.5	41.5	51.0	
	C	.830	.723	.808	
	S.E.	.021	.039	.025	
Judge J	%	53.5	41.0	52.0	50.0
	C	.823	.715	.815	.800
	S.E.	.022	.040	.023	.019

t:t' = 10:200, n = 5, N = 1000

Percent of agreement for average judge = 51.5

Contingency coefficient for average judge = .810 ± .024

TABLE 19
Reliability experiments I and II compared on percentage of agreement

Amount of agreement	Exp. I		Exp. II		Dif.	S.E. dif.	C.R.
	No. of items	% ag.	No. of items	% ag.			
Complete	45	11.0	59	29.5	18.5	3.6	5.15*
Four of five	106	25.9	46	23.0	2.9		
Three of five	126	30.6	53	26.6	10.0		
Two and two	60	14.7	21	10.5	4.2		
Two only	57	13.9	18	9.0	4.9		
None	14	3.4	3	1.5	1.9	1.1	1.75
N	408		200				
Four or better	151	36.9	105	52.5	15.6	4.3	3.62*
Three or better	177	67.5	158	79.0	11.5	3.7	3.10*
For average judge		38.5		51.5	13.0	4.3	3.00*

* P greater than .01

TABLE 20
Reliability experiments I and II compared on percentage of agreement using equal N's*

Amount of agreement	Exp. I		Exp. II		Dif.	S.E. dif.	C.R.
	No. of items	% ag.	No. of items	% ag.			
Complete	12	6.0	59	29.5	23.5	3.6	6.8#
Four of five	62	31.0	46	23.0			
Three of five	66	33.0	53	26.6			
Two and two	30	15.0	21	10.5			
Two only	28	14.0	18	9.0			
None	2	1.0	3	1.5	.5	.9	—
N	200		200				
Four or better	74	37.0	105	52.5	15.5	4.9	3.26#
Three or better	140	70.0	158	79.0	9.0	4.3	2.10

* Same 200 items for both sets of judges.

P greater than .01.

TABLE 21

Categories assigned by judges in reliability experiments I and II: Comparison of proportional use and percentage of agreement

Categories	Amount used: %		Complete agreement: %	
	Exp. I	Exp. II	Exp. I	Exp. II
1. Frustration	6.4	9.6	7.0	10.0
2. Aggression	7.1	8.7	3.0	23.0
3. Submission	8.3	7.1	9.0	14.0
4. Avoidance	8.3	7.0	6.0	43.0
5. Depression	5.2	9.3	15.0	59.0
6. Pleasure	1.1	12.5	28.0	48.0
7. Negative attitudes	6.3	8.7	.0	23.0
8. Positive attitudes	6.6	9.0	.0	33.0
9. Anxiety and conflict	21.8	19.9	15.0	27.0
10. Rationalization	16.0	7.7	9.0	6.0
11. Unclassified	.2	.5	.0	.0
Experiment I				
Ratings per item	5		5	
Number of items	408		200	
N	2040		1000	

TABLE 22

Differences in category frequencies on four forms tested by Chi square

Group	Category Numbers*												Total
	1	3	4	5	6	7	8	9	10	El	Ev	Q	
Men	29	19	40	21	31	41	45	74	42	34	37	34	447
Women	105	91	94	124	134	116	121	234	122	121	111	97	1470
F1	M	18	13	29	15	23	24	50	36	26	24	18	300
	W	49	39	44	50	55	51	45	98	64	60	44	644
F2	M	11	6	11	6	8	17	21	24	6	8	13	147
	W	56	52	50	74	79	65	76	136	58	61	66	826
F1		67	52	73	65	78	75	69	148	100	86	69	944
F2		67	58	61	80	87	82	97	160	64	69	79	973
M	F1	18	13	29	15	23	24	24	50	36	26	24	300
	F2	11	6	11	6	8	17	21	24	6	8	13	147
W	F1	49	39	44	50	55	51	45	98	64	60	45	644
	F2	56	52	50	74	79	65	76	136	58	61	66	826
										X ²	d.f.	P>	
Differences between men's and women's forms													
All										22.986	11	.02	
F1										8.425	11	—	
F2										18.702	11	—	
Differences between Form 1 and Form 2													
All										19.353	11	—	
Men										18.118	11	—	
Women										18.496	11	—	

* See table 4, p. 12, for names of categories.

TABLE 23

Quartile score agreement between Form 1 and Form 2 for subjects writing both forms expressed in contingency coefficients

Subject	Agreement on Single scores		Agreement on Combined scores	
	C	S.E.	C	S.E.
W 1	.630	.075	.655	.095
W 3	.190	.136	.655	.095
W 4	.330	.128	.500	.143
W 5	.450	.113	.655	.094
W 6	.190	.136	.500	.143
W 7	.330	.128	.500	.143
W 8	.190	.136	.280	.182
W 9	.545	.094	.655	.095
W 10	.445	.116	.655	.095
W 11	.040	.134	.280	.182
W 12	.545	.093	.500	.143
W 13	.190	.136	.500	.143
W 14	.040	.134	.500	.143
W 18	.545	.093	.755	.058
M 1	.545	.093	.760	.057
M 2	.040	.128	.655	.085
M 3	.320	.129	.655	.085
M 6	.450	.112	.655	.085
M 7	.532	.093	.760	.057
Average agreement		Percent	C	S.E.
Single scores		41.8	.360	.057
Combined scores		58.0	.600	.050

TABLE 24

Average number of lines per answer written by E and NU groups

Armature Number	E				NU			
	Men		Women		Men		Women	
	N	Mean lines	N	Mean lines	N	Mean lines	N	Mean lines
I	5	4.8	11	6.0	18	6.9	8	6.8
II	5	4.0	12	4.2	10	3.8	8	7.6
III	7	4.3	12	3.3	18	5.4	6	7.0
IV	6	5.5	10	4.6	16	4.7	6	6.2
V	6	4.2	10	3.9	11	5.7	5	4.6
VI	7	5.3	9	4.1	10	4.8	1	2.0
VII	6	4.6	11	3.8	14	6.0	1	3.0
VIII	6	3.5	11	4.7	11	4.7	3	5.0
IX	6	7.2	12	4.2	9	4.5	2	5.5
X	5	3.8	9	4.5	10	4.6	1	4.0
XI	5	3.8	11	4.8	9	4.7	2	6.5
XII	6	7.2	12	4.9	6	3.8	0	—
XIII	6	4.2	9	3.9	16	6.9	3	5.5
XIV	5	4.6	11	4.0	18	7.5	2	5.0
XV	7	4.3	10	3.5	19	5.9	1	3.0
Total		74.3		64.4		79.9		71.7
Average		4.9		4.3		5.3		5.1

TABLE 25
Differences in conflict solution frequencies tested by Chi square

Group	Conflict Solutions*										Total
	At	Re	Ac	Cp	Es	Cf	RF	Ex	Un	No	
F ₂ NU	95	56	69	21	39	7	8	19	3	0	317
E	65	36	62	25	34	12	8	26	24	15	302
F ₁ NU	67	57	58	27	51	13	1	29	3	1	307
F ₂ NU	95	56	69	21	39	7	8	19	3	0	317
M F ₁ NU	22	26	17	6	14	5	0	8	2	1	101
F ₂ NU	28	8	18	8	8	4	1	3	1	0	79
W F ₁ NU	45	31	41	21	37	8	1	21	1	0	206
F ₂ NU	67	48	51	13	31	3	7	16	2	0	238
NU M	50	34	35	14	22	9	1	11	3	1	180
W	112	79	92	34	68	11	8	37	3	0	444
F ₁ NUM	22	26	17	6	14	5	0	8	2	1	101
NUW	45	31	41	21	37	8	1	21	1	0	206
F ₂ NUM	28	8	18	8	8	4	1	3	1	0	79
NUW	67	48	51	13	31	3	7	16	2	0	238
						X ²				d.f.	P>
Differences between NU and E (F ₂)						46.712				9	.01
Differences between Forms 1 and 2											
All						18.332				9	.05
Men						14.444				9	—
Women						17.715				8	.05
Differences between Men and Women											
All						9.879				9	—
F ₁						11.225				9	—
F ₂						12.345				8	—

* See table 4, p. 12.

TABLE 26
Category scores per armature: medians and quartiles* for single score profiles

Category	Whole paper						Armatures I, III, IV, and V		
	Men			Women			Men		
	Q ₁	M	Q ₃	Q ₁	M	Q ₃	Q ₁	M	Q ₃
1.	.00	.82	1.37	.44	.85	1.43	.00	.82	1.31
2.	.00	.00	.00	.00	.00	.00	.00	.00	.00
3.	.00	.49	.99	.12	.62	1.29	.00	.63	1.00
4.	.49	.99	1.74	.38	.72	1.28	.25	.88	1.61
5.	.00	.30	.74	.37	1.05	1.79	.10	.52	.73
6.	.00	.41	.99	1.08	1.55	2.24	.00	.42	1.04
7.	.69	1.33	2.99	.49	1.06	2.24	.59	1.20	2.56
8.	.74	1.19	3.24	.27	.58	1.28	.75	1.31	3.36
9.	2.07	2.49	3.49	1.94	2.77	4.37	1.98	2.55	3.60
10.	.00	.99	1.65	.79	1.19	1.49	.00	.83	1.50
El.	.00	1.24	1.89	1.66	1.31	2.49	.00	1.32	1.72
Ev.	.25	1.24	1.99	.33	1.14	1.99	.36	1.15	2.02
Q.	.00	.49	3.24	.00	.99	1.93	.00	.50	3.16
S	.00	.00	.00	.00	.00	.00	.00	.00	.00
P	.00	.09	.16	.00	.00	.00	.00	.00	.08

* For distributions of 15 NU men and 30 NU women.

TABLE 27

Category scores per armature: medians and quartiles* for combined score profiles

Category	Whole paper						Armatures I, III, IV, and V		
	Men			Women			Men		
	Q ₁	M	Q ₃	Q ₁	M	Q ₃	Q ₁	M	Q ₃
F	7.50	9.10	15.50	8.80	11.34	15.07	7.74	9.01	14.95
C	1.95	3.61	19.20	2.40	4.45	7.20	1.92	3.75	19.60
I	.00	.00	.50	.00	.00	.22	.00	.00	.06
FR	.66	1.70	4.75	2.20	3.06	4.45	.71	1.70	4.85
CR	.00	.95	2.28	2.26	.95	2.95	.00	1.01	2.20
IR	.00	.00	.00	.00	.00	.00	.00	.00	.00
OC	.00	1.00	2.00	.00	.00	.55	.00	1.00	2.00
OA	.00	.00	.00	.00	.00	.00	.00	.00	.00
NA	4.00	5.00	8.00	6.00	9.00	12.00	—	—	—

* For distributions of 15 NU men and 30 NU women.

BIBLIOGRAPHY

1. ALLPORT, G. W. *Personality: a psychological interpretation*. New York: Holt, 1937. Chap. II. Defining personality. pp. 24-54.
2. BALKEN, EVA RUTH, & MASSERMAN, J. H. The language of phantasy. III. The language of phantasies of patients with conversion hysteria, anxiety state and obsessive compulsive neuroses. In Tomkins, Silvan S. *Contemporary Psychopathology*. Cambridge: Harvard Univ. Press. pp. 244-253.
3. BALKEN, EVA RUTH, & VANDERVEER, A. H. The clinical application of a test of imagination to neurotic children. *Amer. J. Orthopsychiat.*, 1942, 12, 68-81.
4. BECK, S. J. Rorschach's test. New York: Grune & Stratton, 1944. Vol. I. Elementary principles.
5. BROWN, J. F. *Psychodynamics of abnormal behavior*. New York: McGraw-Hill, 1940. Chaps. 9-12. Psychoanalysis. pp. 139-256.
6. CHILD, I. S. The use of interview data in quantifying the individual's role in the group. *J. abnorm. soc. Psychol.*, 1943, 38, 301-318.
7. DESPERT, J. LOUISE, & POTTER, H. W. The story: a form of directed phantasy. *Psychiat. Quart.*, 1936, 10, 619-638.
8. FRANK, L. K. Projective methods for the study of personality. *J. Psychol.*, 1939, 8, 389-413.
9. GARRETT, H. E. *Statistics in psychology and education*. New York: Longmans Green, 1940.
10. GRINGS, W. W. The verbal summator technique and abnormal mental states. *J. abnorm. soc. Psychol.*, 1942, 37, 529-45.
11. HARRISON, R. Studies in the use and validity of the Thematic apperception test with mentally disordered patients. II. A quantitative validity study. *Character & Pers.*, 1940, 9, 122-133.
12. HARRISON, R. Studies in the use and validity of the Thematic apperception test with mentally disordered patients. III. Validation by the method of "blind analysis." *Character & Pers.*, 1940, 9, 134-138.
13. HARRISON, R. The Thematic apperception and Rorschach methods of personality investigation in clinical practice. *J. Psychol.*, 1943, 15, 49-74.
14. KRUGMAN, M. Out of the inkwell. *Character & Pers.*, 1940, 9, 91-110.
15. LERNER, E. & MURPHY, L. B., et al. Studying child personality. *Monogr. Soc. Res. Child Develpm.*, 6, No. 4, 1941.
16. LINDQUIST, E. F. *Statistical analysis in educational research*. New York: Houghton Mifflin, 1940.
17. MACFARLANE, JEAN WALKER. Problems of validation inherent in projective methods. *Amer. J. Orthopsychiat.*, 1942, 12, 405-411.
18. MUNROE, RUTH. An experiment in large scale testing by a modification of the Rorschach method. *J. Psychol.*, 1942, 13, 229-263.
19. MURRAY, H. A. Effect of fear on estimations of maliciousness of other personalities. In Tomkins, Silvan S. *Contemporary Psychopathology*. Cambridge: Harvard Univ. Press, 1938, pp. 545-560.
20. MURRAY, H. A. *Explorations in personality*. New York: Oxford Univ. Press, 1938.
21. MURRAY, H. A. *Manual for the Thematic apperception test*. Cambridge: Harvard Univ. Press, 1943.
22. ROSENZWEIG, S. & SHAKOW, D. Play techniques in schizophrenia and other psychoses. II. Schizophrenic constructions. *Amer. J. Orthopsychiat.*, 1937, 7, 36-47.
23. SARGENT, HELEN D. Projective methods: their origins, theory, and application in personality research. *Psychol. Bull.*, forthcoming issue.
24. SEARS, R. R. Experimental studies of projection. I. Attribution of traits. In Tomkins, Silvan S. *Contemporary Psychopathology*. Cambridge: Harvard Univ. Press, 1943. pp. 561-571.
25. TOMKINS, S. S. Limits of material obtainable in the single case study by daily administration of the Thematic apperception test. *Psychol. Bull.*, 1942, 39, 490. Abstract.
26. VERNON, P. E. The matching method applied to investigations of personality. *Psychol. Bull.*, 1936, 33, 149-177.
27. WHITE, R. W. The interpretation of imaginative productions. In Hunt, J. McV. *Personality and the Behavior Disorders*. New York: Ronald Press, 1944. pp. 214-254.
28. YOUNG, K. *Personality and social adjustment*. New York: F. S. Crofts, 1941. Chaps. 19-23, pp. 468-621.